

DOI: 10.19663/j.issn2095-9869.20180624001

http://www.yykxjz.cn/

郜星晨, 姜伟, 刘绍平. 长江中游常见鱼类DNA条形码研究及其电子芯片分析, 2019, 40(5): 89–100

Gao XC, Jiang W, Liu SP. DNA Barcoding and electronic microarrays for common fish species in the middle reaches of the Yangtze River. Progress in Fishery Sciences, 2019, 40(5): 89–100

长江中游常见鱼类 DNA 条形码研究 及其电子芯片分析*

郜星晨¹ 姜伟^{1①} 刘绍平²

1. 中国长江三峡集团有限公司中华鲟研究所 三峡工程鱼类资源保护湖北省重点实验室 宜昌 443100;
2. 农业农村部长江中上游渔业资源环境科学观测实验站 中国水产科学研究院长江水产研究所 武汉 430223)

摘要 为建立长江中游常见鱼类的快速鉴别方法, 文献调研了 7 目 11 科 50 属 64 种鱼类名录, GenBank 共获取 168 条线粒体细胞色素 c 氧化酶 I (CO I) 序列, 分析了序列特征、不同阶元 Kimura-2-paramater (K2P) 遗传距离及系统进化关系。结果显示, 64 种鱼类的种间遗传距离(平均值为 0.084) 明显大于种内(平均值为 0.0079), NJ 树上不同物种均能以较高支持度聚类成独立分支, 以线粒体 CO I 序列作为 DNA 条形码可准确鉴定所研究鱼类; 综合利用分子生物学软件筛选物种特异性探针, 最终 43 种鱼类可筛选出 112 条物种特异性探针, 物种识别率为 67.2%。本研究验证了 DNA 条形码芯片技术在长江中游鱼类物种鉴定的可行性, 可为该地区鱼类物种多样性保护提供技术支持。

关键词 长江中游; DNA 条形码; 电子芯片

中图分类号 S931.1 **文献标识码** A **文章编号** 2095-9869(2019)05-0089-12

长江是世界第三、中国第一大河流, 宜昌至湖口县为中游段(29°30'~30°36'N, 113°24'~115°58'E)。中游江宽流曲, 河网发达, 水质肥沃, 饵料生物丰富(Yi *et al.*, 2011), 与通江湖泊形成了独特而完整的江湖复合生态系统。水流湍急和江河缓流中的鱼类相互衔接和交替, 形成了具有较广种类的江河平原区系复合体鱼类区系(胡军华等, 2008)。中游水域孕育着丰富的鱼类资源, 其中不乏珍稀、名贵以及重要的经济鱼类, 是我国宝贵的淡水鱼类种质资源库。近半个世纪以来, 环境污染、水工建设、围湖造田以及过渡捕捞等造成了鱼类生境片段化、栖息地逐渐丧失, 严重影响

了其生存和繁殖(Yi *et al.*, 2011)。《长江鱼类志》记录中游鱼类 119 种(湖北省水生生物研究所鱼类研究室等, 1976), 根据新的研究, 删除同种异名、偶见种, 整理得到长江中游鱼类 97 种, 隶属 14 目 22 科 68 属。2003~2008 年, 在长江中游及洞庭湖、鄱阳湖共监测到鱼类 14 目 22 科 86 种(胡军华等, 2008); 2009~2010 年的调查结果为 6 目 10 科 55 种(刘春池等, 2012)。可见, 长江中游渔业资源严重衰退, 鱼类种群结构在不断简化, 生态系统功能逐步退化。有效利用自然资源、保护中游地区鱼类多样性成为紧迫而艰巨任务。然而, 鱼类种类繁多, 传统的鱼类分类主要依据外部

* 国家重点研发计划“典型脆弱生态修复与保护研究”重点专项(2016YFC0502210)和农业农村部长江中上游渔业资源环境科学观测站共同资助 [This work was supported by National Key R&D Program of China “Ecological Restoration and Conservation Research of Typical Fragile Ecosystems” (2016YFC0502210), and Scientific Observing and Experimental Station of Fishery Resources and Environment in the Upper and Middle Reaches of Yangtze River, Ministry of Agriculture and Rural Affairs]. 郜星晨, E-mail: 13039829682@126.com

① 通讯作者: 姜伟, E-mail: 106635291@qq.com

收稿日期: 2018-06-24, 收修改稿日期: 2018-07-28

形态特征,不同生长阶段其形态特征的差异给鉴定及分类带来了许多困难(柳淑芳等, 2016)。Hebert 等(2003)提出的 DNA 条形码概念,即通过一段标准的 DNA 序列对物种进行鉴定,为鱼类的快速鉴定提供了新的思路。在加拿大、古巴、北美、阿根廷等国家不同种类的鱼中得到了成功应用(Hubert *et al.*, 2008; Lara *et al.*, 2010; April *et al.*, 2011; Mabragana *et al.*, 2011)。中国也开展了 DNA 条形码的鱼类鉴别研究和应用。彭居俐等(2009)成功鉴定了鲃属(*Erythroculter*)的 4 种鱼, Zhang 等(2012)鉴别了中国南海的 242 种 1336 条鱼类。作者曾经以 DNA 条形码技术实现了鲃形目(Tetraodontiformes)鱼类的准确鉴定,证实 CO I 基因可以用于硬骨鱼类物种鉴定(郜星晨, 2005)。随着分子条形码技术的不断发展,结合电子芯片自身快速、灵敏及高通量等优点, DNA 条形码芯片技术已成为一种识别鉴定物种及确定阶元归属的理想工具。但国内外相关的研究工作相对较少, Kochzius 等(2008)将芯片技术用于鱼类鉴定,证明了 16S rRNA 基因用于设计微阵列探针区分欧洲海洋至少 11 种鱼类的实用性。柳淑芳等(2016)筛选了山东沿海常见的 77 种常见鱼类 DNA 条形码物种特异性探针,最终有

64 种鱼类可用于制作 DNA 条形码芯片。关于我国淡水鱼类的 DNA 芯片适用性研究尚未开展,因此,本研究拟根据长江中游鱼类种类分布情况,探讨 DNA 条形码技术的适用性,并筛选物种特异性探针,从而为长江中游常见鱼类的 DNA 条形码芯片研发工作提供基础资料和技术支撑。

1 材料与方法

1.1 实验材料

根据中国知网数据库(CNKI)查到的 100 余篇主题为“长江中游鱼类”、“鱼类多样性”、“鱼类物种组成”和“渔业资源调查”的调查文献和相关科学考察集,并参考往年监测记录作补充(胡军华等, 2008; Yi *et al.*, 2011; 刘春池等, 2012),确定长江中游常见鱼类物种名录,即 7 目 11 科 50 属 64 种鱼类。物种的分类和名称通过 Fishbase 数据库(www.fishbase.org)、《中国动物志》(乐佩琦, 2000)及分类文献确认。从 GenBank 下载对应物种收录的 168 条 CO I 序列(表 1),尽量选取研究区域内不同的地理种群,每条序列需经 BOLD 数据库验证。

表 1 本研究中 64 种鱼类 CO I 基本信息
Tab.1 Information of CO I genes of 64 species

分类 Taxonomic status	物种 Species	尾数 Quantity	GenBank 序列号 GenBank Accession No.
鲃形目 Acipenseriformes 鲃科 Acipenseridae	中华鲃 <i>Acipenser sinensis</i>	2	KM366079, KM366080
鲑形目 Salmoniformes 银鱼科 Salangidae	太湖新银鱼 <i>Neosalanx taihuensis</i>	2	JX524196, NC_019629
颌针鱼目 Beloniforme 鱻科 Hemiramphidae	间下鱻 <i>Hyporhamphus intermedius</i>	2	KP260625, NC_026467
鲱形目 Clupeiformes 鳊科 Engraulidae	鳊 <i>Coilia ectenes</i>	2	HM180534-5
鲤形目 Cypriniformes 鲤科 Cyprinidae	华鲮 <i>Sinilabeo rendahli</i>	2	KT336714, NC_028169
	墨头鱼 <i>Garra pingi</i>	2	HM536884, JX290078
	胭脂鱼 <i>Myxocyprinus asiaticus</i>	2	AY986503, AP006764
	厚颌鲂 <i>Megalobrama pellegrini</i>	2	JX403698, JX242529
	泉水鱼 <i>Pseudogyrinocheilus prochilus</i>	2	GU086572, JX074193
	岩原鲤 <i>Procypris rabaudi</i>	2	EU082030, KR862062
	鲤 <i>Cyprinus carpio</i>	6	KR861871, HQ600721-5
	鲫 <i>Carassius auratus</i>	2	KR861871-2
	鲢 <i>Hypophthalmichthys molitrix</i>	3	KR862041, KR862047-8
	鳙 <i>Hypophthalmichthys nobilis</i>	2	KR862052-53
	红鳍鲃 <i>Culter erythropterus</i>	2	HQ536351-2
	贝氏鲃 <i>Hemiculter bleekeri</i>	2	KR861975, KR861957
	银飘鱼 <i>Pseudolaubuca sinensis</i>	2	KR862140-1

续表 1

分类 Taxonomic status	物种 Species	尾数 Quantity	GenBank 序列号 GenBank Accession No.
鲤形目 Cypriniformes	寡鳞银飘鱼 <i>Pseudolaubuca engraulis</i>	2	KR862092-3
鲤科 Cyprinidae	鳊 <i>Parabramis pekinensis</i>	2	KR862056-7
	草鱼 <i>Ctenopharyngodon idellus</i>	2	KR861832-3
	鳊 <i>Elopichthys bambusa</i>	4	KR861908-11
	赤眼鳟 <i>Squaliobarbus curriculus</i>	2	KR862252-3
	马口鱼 <i>Opsariichthys uncirostris</i>	2	KR862054, DQ367044
	宽鳍鱮 <i>Zacco platypus</i>	3	HQ536624-5, HQ536601
	大鳍鱮 <i>Acheilognathus macropterus</i>	3	KR861750-52
	兴凯鱮 <i>Acheilognathus chankaensis</i>	2	NC_023101, KF695385
	高体鲮 <i>Rhodeus ocellatus</i>	3	KM610905-6, KR862163
	中华鲮 <i>Rhodeus sinensis</i>	6	HQ536481-6
	细鳞斜颌鲴 <i>Xenocypris microlepis</i>	3	GU434334-6
	银鲴 <i>Xenocypris argentea</i>	3	KR862281-2, KR862269
	似鳊 <i>Pseudobrama simoni</i>	4	KR862083-6
	花鲮 <i>Hemibarbus maculatus</i>	3	KR861927-9
	唇鲮 <i>Hemibarbus labeo</i>	3	HQ536370-2
	蛇鲴 <i>Saurogobio dabryi</i>	3	KR862172-73, HQ536513
	圆筒吻鲴 <i>Rhinogobio cylindricus</i>	2	KF724884, NC_024540
	长鳍吻鲴 <i>Rhinogobio ventralis</i>	2	KR862162, KR862158
	吻鲴 <i>Rhinogobio typus</i>	2	KR862142-43
	似刺鳊鲴 <i>Paracanthobrama guichenoti</i>	2	KJ645748, NC_024430
	圆口铜鱼 <i>Coreius guichenoti</i>	3	KP054470-72
	铜鱼 <i>Coreius heterodon</i>	5	KR861813-7
	黑鳍鲈 <i>Sarcocheilichthys nigripinnis</i>	3	HQ536510-12
	华鲈 <i>Sarcocheilichthys sinensis</i>	5	KR862164-6, KR862168-9
	银鲴 <i>Squalidus argentatus</i>	3	KR862239-41
	麦穗鱼 <i>Pseudorasbora parva</i>	3	HQ961065-6, KC688868
	棒花鱼 <i>Abbottina rivularis</i>	3	KR861723-4, KR861747
	宜昌鳅鲃 <i>Gobiobotia filifer</i>	2	KP325413, NC 029187
鲤形目 Cypriniformes	犁头鳅 <i>Lepturichthys fimbriata</i>	2	JN177229, JN177162
平鳍鳅科 Balitoridae	中华金沙鳅 <i>Jinshaia sinensis</i>	2	JN177114, JX155733
	壮体沙鳅 <i>Sinibotia robusta</i>	2	KC871170, JN177235
	花斑副沙鳅 <i>Parabotia fasciata</i>	4	KC871165-68
	紫薄鳅 <i>Leptobotia taeniops</i>	2	JN177236, JN177224
	长薄鳅 <i>Leptobotia elongata</i>	2	JN177225, JX155734
鲈形目 Perciformes	鳊 <i>Siniperca chuatsi</i>	3	EF143387, KP112454-55
鲈科 Serranidae	斑鳊 <i>Siniperca scherzeri</i>	2	AP014527, JQ010986
鲈形目 Perciformes	子陵吻鰕虎鱼 <i>Rhinogobius giurinus</i>	3	KM610888-91
鰕虎鱼科 Gobiidae	褐吻鰕虎鱼 <i>Rhinogobius brunneus</i>	4	HQ536455-58
鲈形目 Perciformes	乌鳢 <i>Channa asiatica</i>	2	KC819604-5
鳢科 Channidae			
鲇形目 Siluriformes	长吻鲇 <i>Leiocassis longirostris</i>	1	NC_030628
鲇科 Bagridae	粗唇鲇 <i>Leiocassis crassilabris</i>	2	KC768227, JX867257
	凹尾拟鲇 <i>Pseudobagrus emarginatus</i>	2	NC_024279, KJ600628
	大鳍鲇 <i>Hemibagrus macropterus</i>	3	JF292350-52
	黄颡鱼 <i>Tachysurus fulvidraco</i>	4	KP112370-73
鲇形目 Siluriformes	胡子鲇 <i>Silurus asotus</i>	2	JX087351, JN116720
鲇科 Siluridae	中华纹胸鲃 <i>Glyptothorax hainanensis</i>	2	DQ514360, HQ593580

1.2 实验方法

1.2.1 条形码分析 Bioedit 比对切齐序列, MEGA 6.0 计算序列的序列变异、碱基组成及基于 Kimura-2-paramater(K2P)模型不同分类阶元的遗传距离(Tamura *et al.*, 2013)。进一步构建邻接(Neighbour-Joining, NJ)系统进化树, 将 Bootstrap 值参数设为 1000。Microsoft Excel 2007 作遗传距离频率分布直方图。

1.2.2 DNA 条形码芯片探针设计和分析 建立本地 Blast 数据库, 序列文件 barcode.fasta 格式化, 形成文件 barcode.fasta.nsq。由于序列数目较多, 为了提高筛选效率, 鲤形目 48 种鱼类 130 条序列划分为一个数据集, 其余 16 种鱼类 38 条序列为另一数据集。各数据集分别利用 OligoArray 2.1 筛选探针, 探针长度为 23~27 bp, T_m 值为 70°C~82°C, GC 含量为 40%~55%, Na^+ 和 DNA 浓度分别为 1 mol/L 和 1 μ mol/L, 形成二级结构的最大 T_m 值和交叉杂交最小 T_m 值为 65°C, 两探针间距离为 25 bp, 最大探针数为 10, 连续出现的单一碱基数不超过 4 个。

选取该物种序列所共有且与其他物种序列无同

源性的探针作为该物种的探针, 在线软件 OligoCalc (<http://www.simgene.com/OligoCalc>) 优化, 排除易形成发夹、茎环及自身二聚体结构的不合格探针。整理各数据集筛选合格的探针与切齐序列于在线软件 OHM(<http://bioinfo.unice.fr/software/ohm/>) 中进行虚拟杂交, 查看杂交结果。

2 结果与分析

2.1 DNA 条形码分析结果

2.1.1 DNA 条形码序列特征 经同源比对得到长度为 624 bp 的有效片段, 其中, T=28.7%, C=28.2%, A=25.3%, G=17.8%(表 2)。A+T 含量为 54.0%, 第 1、2 和 3 密码子的 G+C 含量分别为 57.5%、42.7%和 37.6%, 表现出明显的反 G 偏倚, 与其他硬骨鱼类线粒体的特点一致。在碱基变异中, 不变位点为 516 个, 占总位点数的 82.7%。转换位点为 66 个, 颠换位点为 41 个, 86.4%的碱基转换和 97.6%的碱基颠换发生在第 3 密码子位点。转换与颠换比(R)为 3.58, 表明此片段没有饱和, 适合进行系统发育分析。

表 2 64 种鱼类 CO I 基因序列组成统计
Tab.2 The content of CO I gene sequences of 64 species

位点 Sites	L	ii	S_i	S_v	T	C	A	G	A+T	$R=S_i/S_v$
全部位点 Total sites	624	516	66	41	28.7	28.2	25.3	17.8	54.0	3.58
第 1 位点 Codon 1 point	208	198	8	1	17	26.9	25.3	30.6	42.3	6.63
第 2 位点 Codon 2 point	208	207	1	0	42	28.8	15.4	13.9	57.4	2.46
第 3 位点 Codon 3 point	208	110	57	40	27	28.8	35.3	8.8	62.3	1.42

注: L: 长度; ii: 不变位点; S_i : 转换位点; S_v : 颠换位点

注: L: Length; ii: Invariant site; S_i : Conversion site; S_v : Transversion site

2.1.2 64 种鱼类 K2P 遗传距离分析 本研究中鱼类 K2P 遗传差异变化幅度较大, 0~0.298 不等。种内遗传差异在 0~0.0776 之间(*P.prochilus*), 均值为 0.0079。59 种(占总类数 92.1%) 鱼类种内差异小于 2%, 其余 5 种鱼类, 即马口鱼(*O.uncirostris*)、乌鳢(*C.asiatica*)、斑鳅(*S.scherzeri*)、泉水鱼(*P.prochilus*) 等遗传距离大于 2%(表 3 和表 4)。属内种间遗传差异为 0.003 (*Rhinogobio*)~0.19 (*Rhodeinae*), 均值为 0.084, 该范围与种内遗传差异的分布范围重叠(图 1)。其中, 圆筒吻鲷(*R.cylindricus*) 与 吻鲷(*R.typos*) 间为 0.003, 长吻鲷(*L.longirostris*) 和 粗唇鲷(*L.crassilabris*) 间为 0.01, 花鳍(*H.maculatushe*) 和 唇鳍(*H.labeo*) 间为 0.02, 大鳍鲷(*A.macropterus*) 与 兴凯鲷(*A.chankaensis*) 间为 0.035, 鳅(*S.chuatsi*) 和 斑鳅(*S.scherzeri*) 间为 0.048, 细鳞斜颌鲷(*X.microlepis*) 和 银鲷(*X.argentea*) 间为 0.065, 铜鱼(*C.heterodon*) 和 圆口铜鱼(*C.guichenoti*) 间为 0.075, 这些

鱼类属内种间遗传距离均小于属内种间平均遗传差异。同科属间遗传差异从 0.018 (*Balitoridae*)~0.248 (*Balitoridae*), 均值为 0.184。目内科间遗传差异从 0.204 (*Cypriniformes*)~0.288 (*Perciformes*), 均值为 0.232。目间遗传差异从 0.227(*Perciformes* 与 *Beloniforme*)~0.298 (*Acipenseriformes* 与 *Beloniforme*) 间, 均值为 0.253。

2.1.3 64 种鱼类系统进化树 本研究中的 64 种鱼类不同个体均能聚类形成种的单系支, 且支持率较高, 50 个属不同物种也能聚类成属的单系支。但是, 科内属间、目内科间及目间聚类效果不明显, 不能形成平行的独立分支, 与传统分类存在一些分歧(图 2)。

2.2 DNA 条形码芯片分析结果

2.2.1 数据集 1 探针信息及杂交结果 初步筛选出 46 个物种的 130 条探针, OligoCalc 优化后去除 12 条不合格探针, 获得 43 个物种的 118 条探针(表 3)。所得探

表 3 数据集 1 探针编号及序列
Tab.3 Probe IDs and sequences of dataset 1

探针编号 Probe ID	探针序列 Probe sequence	中文名(拉丁名) Species (Latin name)	种内遗传距离 Genetic distance within species
TZ1	CTGACTCGTACCCCTAATGATTG	鲢 <i>H.molitrix</i>	0.43
TZ2	TTCTATCCCTACCCGTGCTAGCTG	银飘鱼 <i>P.sinensis</i>	0.16
TZ3	CTGCTACTAGCCTCTTCTGGCGTT		
TZ4	AGACCGAAACCTAAATACTACCTTC	寡鳞银飘鱼 <i>P.engraulis</i>	0
TZ5	GTGTGTCCTCAATTTTAGGCGCAAT		
TZ6	GCCGGAATTACAATACTCCTAAC	草鱼 <i>C.idellus</i>	0
TZ7	TCCTTCTCTATCTCTACCAGTTC		
TZ8	TTGGAGGGTTAGGTAAGGGGCTT	马口鱼 <i>O.uncirostris</i>	2.28
TZ9	CCACTTAGCAGGAGTTTCATCAA		
TZ10	CTGCTACTATCTCTGCCAGTCT	高体鳊 <i>R.ocellatus</i>	1.97
TZ11	TGTTTGAGCTGTACTCGTTACAG		
TZ12	CGATCTTTTCACTACACCTGGCA		
TZ13	ACTTTTACTAGCCTCCTCTGGGG		
TZ14	GACTTGTTCCTCTTATAATTGGGG		
TZ15	TACTGCTGTCTCTGCCGGTCTTA	中华鳊 <i>R.sinensis</i>	0.21
TZ16	TTCTTGCCTCCTCCGGAGTTGAA		
TZ17	CCTTACAGACCGAAATCTTAACAC	银鲴 <i>X.argentea</i>	0.21
TZ18	TTCTCCTATCACTACCAGTGCTAG		
TZ19	CCGATCCTGTATCAACACCTATTC	细鳞斜颌鲴 <i>X.microlepis</i>	0.32
TZ20	CCTTCTCCTATCACTACCTGTTCTA	鲫 <i>C.auratus</i>	0.32
TZ21	TTTGTGGATCCGTAATTGTAACCG		
TZ22	TTCTGTACTACTAGCTTCCTCTG		
TZ23	AAGTACTTGTACCCCTGATAATCG		
TZ24	CGTATGAGCTGTACTCGTAACAG	花斑副沙鳅 <i>P.fasciata</i>	0
TZ25	GTGTATCATCCATTCTTGAGCA		
TZ26	GCACATGCCTTTGTCATAATCTTC		
TZ27	ACCATCTTCTCCCTCCACCTAGCA	长薄鳅 <i>L.elongata</i>	0.16
TZ28	TATCCTTATCGGAGGCTTCGGGAA		
TZ29	GCCGTACTCGTAACAGCCGTTTT	紫薄鳅 <i>L.taeniops</i>	0.16
TZ30	ATACCAGACACCCCTCTTCGTTT		
TZ31	TCTAGCAGGTGTATCATCCATCCT		
TZ32	CCGTAGACCTTACCATCTTCTCCC		
TZ33	ATCCTTATTGGAGGCTTTGGGAACT		
TZ34	ATTTATAATGTCATCGTCACTGCGC		
TZ35	CTGGAATCACAATGCTTCTAACG	兴凯鳊 <i>A.chankaensis</i>	0
TZ36	CGGACCGTAATCTAAACACTACA	大鳍鳊 <i>A.macropterus</i>	0.48
TZ37	ACTTCTTCTTATCTCTACCAGTC	厚颌鲂 <i>M.pellegrini</i>	0
TZ38	ACCGATCGAAATCTTAACACCAC	墨头鱼 <i>G.pingi</i>	0.48
TZ39	TCTATTATCACTACCAGTGCTAGC		
TZ40	GATGACCAAATCTACAACGTTATCG		
TZ41	TTCTCTCTACCCGTTTTAGCTG	鳊 <i>E.bambusa</i>	0.11
TZ42	GACCCCATTTTATATCAGCACCT		
TZ43	CTTCTACTCCTAGCTTCTTCTGG	鲤 <i>C.carpio</i>	0.21
TZ44	ATGAACCGTATACCCACCTCTTG		

续表 3

探针编号 Probe ID	探针序列 Probe sequence	中文名(拉丁名) Species (Latin name)	种内遗传距离 Genetic distance within species
TZ45	GATTACTATACTCCTTACAGACCGA	红鳍鲌 <i>C.erythropterus</i>	0
TZ46	GGAATTACTATGCTCCTCACAGA	鳊 <i>P.pekinensis</i>	0.32
TZ47	TCCCTCTTTCCTTCTATTACTTGC		
TZ48	CCAATCCTATATCAGCACCTGTT		
TZ49	TTCTCCTGTCCTTACCCGTA CTG	宽鳍鱲 <i>Z.platyus</i>	0.21
TZ50	TGTTACTAGCCTCTTCAGGTGTT		
TZ51	CTTATTCGTCTGAGCCGTA CTG		
TZ52	GGGTTTCATCAATCTTAGGAGCA		
TZ53	CTGCCGTA CTACTTCTTCTATCA	华鲮 <i>S.rendahli</i>	0
TZ54	TTAGGGGCCATCAACTTCATTAC		
TZ55	CTATTACTACTTGCCTCTTCTGGT		
TZ56	CTTCTCTACTCTCCCTACCAGTC	犁头鳅 <i>L.fimbriata</i>	0
TZ57	AAACACAACCTTCTTCGACCCAG	壮体沙鳅 <i>S.robusta</i>	0
TZ58	ACTTCTATCTTTGCCCGTCTTAG		
TZ59	GCGCATCTGTTGATTTAACCATC		
TZ60	GTATTACAATGCTCCTCACAGAC	贝氏鲮 <i>H.bleekeri</i>	0.16
TZ61	TAATACTACCTTCTTTGACCCGG	麦穗鱼 <i>P.parva</i>	0
TZ66	GCTTAGTCAACCAGGATCACTCCT	宜昌鳅鲇 <i>G.filifer</i>	0
TZ67	GGAAGTACTTGTGCCATTCATGA		
TZ68	GACAGGCTGAACAGTTTATCCTCC		
TZ69	ACGATTAATATAGAACCTCCAGCC		
TZ70	TACTTCTTCTGCTTTCGCTGCCCG		
TZ71	TACTTCTGTCCCTACCCGTCCTA	泉水鱼 <i>P.prochilus</i>	7.76
TZ72	CGATCCTTGGGGCAATCAATTTT		
TZ73	TCATTCTGCTTCTCCTAGCCTCTT		
TZ74	CGACGACCAAATCTACAATGTTATC		
TZ75	CGATTCTTTACCAACACCTGTTC		
TZ76	GCATTTCCCGGATAAAACAACAT	胭脂鱼 <i>M.asiaticus</i>	0
TZ77	CGGAGCCTCTGTAGATTTAACTA		
TZ78	GACAGACCGAAACCTAAACACAA		
TZ79	CTGGGCGACGATCAAATTTATAATG	似鳊 <i>P.simoni</i>	0.08
TZ80	TCCTCCTGCTACTTGCTTCTTCTGG		
TZ81	CGCTTCACTTAGCAGGTGTATCATC		
TZ82	CAGACCGAAACCTTAACACAACA		
TZ83	GCTCCTAGGAGATGACCAAATTTA	花鲢 <i>H.maculatus</i>	0.11
TZ84	CCGGAATTACAATACTCCTCACAG		
TZ85	TTTCTACTTCTGCTCGCCTCTTCT		
TZ86	TCCTACTCCTGCTAGCCTCTTCT	唇鲮 <i>H.labeo</i>	0.32
TZ87	ACTGGTTCCTAATGATTGGAG	蛇鮈 <i>S.dabryi</i>	5.23
TZ88	GTGTTTCATCGATTTTAGGCGCT		
TZ89	CTCCTTCTCCTATCACTGCCTGT		
TZ90	TAGCCGCTGGAATTACAATACTC		
TZ91	TAATACCACCTTCTTTGACCCAG		
TZ92	TCACTACCACGATCAACATGAAG		

续表 3

探针编号 Probe ID	探针序列 Probe sequence	中文名(拉丁名) Species (Latin name)	种内遗传距离 Genetic distance within species
TZ93	TCTTCTCCTACTGGCCTCTTCTG	吻鲟 <i>R.typus</i>	0
TZ94	CTCTTTGTTTGAGCTGTGCTTGT		
TZ95	CGTACTTCTCCTTTTATCACTGC		
TZ96	CCCCTATTTGTCTGATCTGTACT	圆筒吻鲟 <i>R.cylindricus</i>	0
TZ97	CGTACTTCTCCTGCTATCACTAC		
TZ98	CCGCAGGGATTACTATACTTCTA		
TZ99	TCTCAACTACATTCTTCGACC		
TZ100	TATTAGCCTCTTCCGGAGTCGAG	似刺鲃 <i>P.guichenoti</i>	0
TZ101	CTACTAGCTTCTTCTGGAGTTGAA	铜鱼 <i>C.heterodon</i>	0
TZ102	AATAATCGGAGCACCTGATATAGC		
TZ103	TCCCAATACCAAACCCCTCTTTG	圆口铜鱼 <i>C.guichenoti</i>	0
TZ104	TCTATTACTAGCCTCTTCCGGGGTT		
TZ105	ACTTGTCCCACTAATAATCGGAG	银鲟 <i>S.argentatus</i>	0.32
TZ106	TGCTTCTCCTAGCCTCTTCTGGT		
TZ107	AATAATCGGAGCGCCAGACATGG		
TZ108	ATCAGACACCTTTATTTCGTATGGGCT	华鲟 <i>S.sinensis</i>	0.48
TZ109	CGATCCTGTACCAACTTATTC		
TZ110	CAGACCGAAACCTAAACTACA	黑鳍鲟 <i>S.nigripinnis</i>	0.11
TZ111	CCTAACAATCTTTTCACTTACCTG		
TZ112	TGACTTGTACCCCTTATAATTGGAG		
TZ113	GGCATTCTAGGCGATGATCAAA	棒花鱼 <i>A.rivularis</i>	0.43
TZ114	TCCCACTCATAATTGGAGCACCT		
TZ115	CCCTCGAATGAACAACATAAGCT		
TZ116	AATCTTCTCACTGCACTTAGCGG		
TZ117	GGCCCTCTCCCAATATCAAACAC		
TZ118	TTCTACTTCTCTCGCTGCCCGTT		

注: 加粗为无效探针

Note: The probes in bold are not available

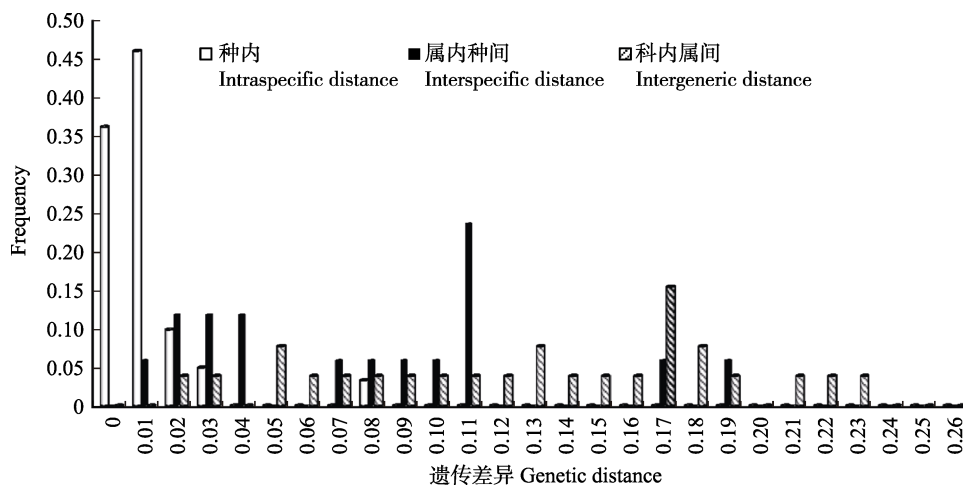


图 1 不同分类阶元水平的遗传差异分布

Fig.1 Frequency distribution of K2P genetic distance within various taxonomic levels

表4 数据集2 探针编号及序列
Tab.4 Probe ids and sequences of dataset 2

探针编号 Probe ID	探针序列 Probe sequence	物种 Species	种内遗传距离 Genetic distance within species
TZ119	CCATCCTCTACCAACACCTATTC	中华鲟 <i>A.sinensis</i>	0.16
TZ120	CGGGGATCACAATACTTCTAACAG		
TZ121	ACCAGACACCTCTATTTGTATGATC		
TZ122	GGTATCGTCCATTTTAGGAGCTATT		
TZ123	TCGTATGAACAATATGAGCTTCTGG		
TZ124	TCATAATTGGCGGATTTCGGAAACTG		
TZ125	TGCTCACAGATCGAAACCTCAAC	乌鳢 <i>C.asiatica</i>	2.79
TZ126	CATTTCCCAATATCAAACACCCCTA		
TZ127	CTATCTCTTCCAGTCTTAGCTGC		
TZ128	ATCCTAATCACCGCCGTACTACTAC		
TZ129	ACTCTCCCTACCAGTACTAGCTGCTG	大鳍鱠 <i>H.macropterus</i>	0.21
TZ130	ACTACTAGCCTCATCTGGTGTTGAA		
TZ131	GCAACCTTAACACAACCTTCTTC	黄颡鱼 <i>T.fulvidraco</i>	0.08
TZ132	CTACTAGCCTCCTCCGGAGTTGAA		
TZ133	ATGAGCTTCTGATTACTACCTCCC		
TZ134	TACTTCTTTCCCTGCCAGTCTGG	胡子鲇 <i>S.asotus</i>	0
TZ135	TTTTTCATTACATCTCGCAGGGGTG		
TZ136	CTCCCTCCCTCCTTCCTACTACTA		
TZ137	GGAAACTGACTTGTACCCCTCATG		
TZ138	TCCTGATTACTGCTGTACTTCTTC	鳊 <i>S.chuatsi</i>	0
TZ139	ATGAGCTTCTGACTACTCCCACCG		
TZ140	ATGCCAATTATAATCGGAGGATTTCG		
TZ141	GTGATCCAATCCTTTACCAACAC	斑鳊 <i>S.scherzeri</i>	2.29
TZ142	TCTGATCACTGCTGTGCTTCTTC		
TZ143	TACGGACCGCAACTTAAACACAA	子陵吻鰕虎鱼 <i>R.giurinus</i>	1.09
TZ144	TCTAATTACAGCTGTCCTGCTAC		
TZ145	GTATCTCATCTATTCTTGGGGCA		
TZ146	CCATCAACTTCATTACGACTATCC	褐吻鰕虎鱼 <i>R.brunneus</i>	1.50
TZ147	TGACTTATTCCACTAATGATCGGTG		
TZ148	TAATACGACCTTCTTTGACCCAG		
TZ149	GTGCTTCTACTATTGTCCCTACC	凹尾拟鲮 <i>P.emarginatus</i>	0
TZ150	GTTATTATCTTCTTCCGGGGTTGA		
TZ151	GACTACTTCCACCCTTTTCCTAT		
TZ152	TCCTTACAGACCGAAACCTAAAC	鲚 <i>C.ectenes</i>	0.16
TZ153	GTTTTAGCTGCCGAATCACAAT		
TZ154	GTATTACCATACTCCTAACAGACCG	粗唇鲃 <i>L.crassilabri</i>	0.16
TZ155	GTGACCCAATTCTATACCAACAC	中华纹胸鲃 <i>G.hainanensis</i>	1.96
TZ156	CCATCTTGTACCAACACCAGTTC	太湖新银鱼 <i>N.taihuensis</i>	0
TZ157	ACGGACCGAAACTTAAACACCAC		
TZ158	AGCTGCAGGCATCACTATGCTTC		
TZ159	CCTTACCGACCGAAACCTTAACA	间下鲃 <i>H.intermedius</i>	0
TZ160	TCCTTGCCGCTGGTATTACTATA		
TZ161	CGTCTACTTCTCCTTTCCCTTC		
TZ162	CGGGTATTTCTTCTATCCTTGGAG		
TZ163	CAGTTGATTTAACGATCTTCTCCC		
TZ164	GGCTCACGCCTTCGTAATAATTTT		

注: 加粗为无效探针

Note: The probes in bold are not available

靶标序列杂交,可作为物种的鉴别探针。

2.2.2 数据集 2 探针信息及杂交结果 初步筛选出 16 个物种的 80 条探针, OligoCalc 优化去除 34 条不合格探针, 得到 15 个物种的 46 条探针(表 4)。探针与对应物种序列在 OHM 中虚拟杂交, 结果显示, TZ154、TZ155、TZ162 分别为粗唇鲮、中华纹胸鲃(*G.hainanensis*)和间下鱊(*H.intermedius*)的探针, 但不能与其靶标序列杂交; TZ120、TZ124~128、TZ135、TZ141~143 和 TZ147 在与其靶标序列结合的同时还

能与非靶标序列结合。最后获得 11 个物种的 32 条探针, 可作为物种的鉴别探针。

总之, 本研究下载的 64 个物种的 168 条 CO I 序列, 利用 OligoArray 2.1 软件初步筛选出 62 个物种的 210 条特异性探针, OligoCalc 优化后选出 58 个种的 164 条探针, OHM 中虚拟杂交后 43 个物种的 112 条探针可作为物种的鉴别探针, 可鉴定的物种数占总物种数的 67.2%(表 5)。

表 5 筛选探针及杂交情况
Tab.5 Probe selection and hybridization conditions

数据集 Dataset	物种数 Quantity of species	初选物种数 Species through first selection	次选物种数 Species through second selection	可鉴定种数 Species can be identified	可鉴定物种数比例 Percentage of identifiable species (%)
1	48	46	43	32	66.7
2	16	16	15	11	68.8
总计 Total	64	62	58	43	67.2

3 讨论

3.1 DNA 条形码的适用性

DNA 条形码是一项利用标准序列快速和准确鉴定的诊断技术, 为分类学发展的新方向(Dayrat, 2005)。Hebert 等(2003)对大量动物的线粒体 CO I 序列分析发现, 2%是物种划分的参考值, 且种间距离至少是种内距离的 10 倍。本研究中, 59 种鱼类(占总类数的 92.1%)种内差异小于 2%, 种内平均遗传距离为 0.0079, 属内种间平均遗传距离为 0.084, 种间距离为种内距离的 11 倍, 基本符合“10 倍种内遗传距离”。构建的系统进化树中, 64 种鱼类均能形成种的单系支, 且支持率较高。这些都充分证明了线粒体 CO I 序列的分子条形码对于长江中游常见鱼类有很好的分类效果, 可进一步筛选 DNA 条形码芯片(Blanchet, 2012; Kim *et al*, 2011)。其中, 49 种(占总类数的 76.6%)鱼类的种内遗传距离小于种间最小遗传距离(0.003), 可形成明显的条形码间隙。其余种类种内与种间遗传差异存在重叠, 未形成 DNA 条形码间隙, 造成这种现象的原因可能是物种快速形成、渐渗杂交或不完全世系分选等(Zhang *et al*, 2012)。此外, 本研究中, 种内、属内种间、科内属间、目内科间及目间的遗传差异分别为 0.0079、0.084、0.184、0.232 和 0.253, 分类单元越高, 遗传距离越大, 为传统分类学的有效性提供了依据。但除了种间与种内距离差异明显外, 物种以上阶元的增加明显变缓慢, 且不同高阶分类单元遗传差异出现重叠, 未形成明显的

DNA 条形码间隙。表明基于线粒体 CO I 序列的 DNA 条形码序列解析能力有限, 可用于物种的分类鉴定, 但不适合于高阶元分类单元的系统进化分析(Dayrat, 2005)。

3.2 电子芯片的适用性

DNA 条形码芯片是一种利用 DNA 条形码序列快速鉴定物种的技术, 可用于构建物种分子监测网络, 快速分析物种群落的演变(Hajibabaei *et al*, 2007)。本研究从 64 种鱼类中初步筛选得到 62 个种的 210 条物种特异性探针, 2 个物种没有筛选出探针。利用 OligoCalc 对筛选出的探针进行检验, 4 个种的探针不合格, 最后获得 58 个物种的 164 条探针。OHM 软件虚拟杂交后, 仅 43 个物种的 112 条探针可以与其靶标序列特异性杂交。种内遗传距离小于 2%的马口鱼、乌鳢、斑鳊、泉水鱼未筛选出物种特异性探针。其中, 圆筒吻鲈、吻鲈、长吻鲮、粗唇鲮、花、唇鲮、大鳍鱮、兴凯鱮、鳊、斑鳊、细鳞斜颌鲷、银鲷、铜鱼、圆口铜鱼等属内种间遗传距离小于平均属内种间遗传距离的物种, 仅银鲷、鳊、铜鱼、圆口铜鱼等获得物种特异性探针。OligoCalc 优化后物种识别率为 90.6%(58/64), OHM 虚拟杂交后物种识别率仅为 67.2%(43/64)。结果表明, 尽管 DNA 条形码基因可以获得大量的物种特异性探针, 但大部分探针存在交叉杂交或缺少杂交信号, 这与 Kochzius 等(2010)对 50 种鱼类的条形码芯片研究结果一致。本研究物种识别率(67.2%)与欧阳小艳等(2007)、柳淑芳等(2016)的结

果相比敏感性较低,这可能与选定物种间亲缘关系有一定关系,序列相似度高,序列的特异性片段就少。此外,624 bp 的 *CO I* 序列包含信息有限,尽管能够有效区分物种,但似乎无法满足基于该序列筛选的探针用于物种鉴别的需求。综上所述,本研究在理论上验证了 DNA 条形码长江中游鱼类鉴定中的可行性,筛选的物种特异性探针具有较好的鉴定精度,为下一步电子芯片的制备与应用奠定了基础。考虑探针的计算机模拟的性能和实际的实验杂交之间存在差异,一方面应加强探针杂交的动力学过程研究,提高预测探针的性能,并增加杂交信号稳定性;另一方面应在实际操作中验证探针的物种识别能力,进一步优化获得的高质量分子探针。此外,由于用于本研究分析的数据有限,未知的单倍型和地理变异种会造成无法检测或错误鉴定(柳淑芳等,2016),今后的研究中应增加各个地理分布区物种的条形码序列。

参 考 文 献

- April J, Mayden RL, Hanner R, *et al.* Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the USA*, 2011, 108(26): 10602–10607
- Blanchet S. The use of molecular tools in invasion biology: An emphasis on freshwater ecosystems. *Fisheries Management and Ecology*, 2012, 19(2): 120–132
- Dayrat B. Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 2005, 85(3): 407–415
- Fish Research Laboratory of Institute of Hydrobiology. *Fish of the Yangtze River*. Beijing: Science Press, 1976 [湖北省水生生物研究所鱼类研究室. 长江鱼类. 北京: 科学出版社, 1976]
- Gao XC. DNA barcoding of Tetraodontiformes in coastal waters of China. Master's Thesis of Jinan University, 2015 [郜星辰. 基于线粒体 *CO I* 序列的中国沿海鲀形目鱼类 DNA 条形码研究. 暨南大学硕士研究生学位论文, 2015]
- Hajibabaei M, Singer GA, Clare EL, *et al.* Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biology*, 2007, 5(1): 24–24
- Hebert N, Cywinska A, Ball SL, *et al.* Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 2003, 270(1512): 313–321
- Hu JH, Zhang CL, Hu HJ. Structure and variation of fish diversity of lake in the central Yangtze River. *Journal of Hydroecology*, 2008, 1(1): 47–51 [胡军华, 张春兰, 胡慧建. 长江中游湖泊鱼类物种多样性结构及动态. 水生态学杂志, 2008, 1(1): 47–51]
- Hubert N, Hanner R, Holm E, *et al.* Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One*, 2008, 3(6): e2490
- Kim S, Koo H, Kim J, *et al.* DNA chip for species identification of Korean freshwater fish: A case study. *Biochip Journal*, 2011, 5(1): 72–77
- Kochzius M, Nolte M, Weber H, *et al.* DNA microarrays for identifying fishes. *Marine Biotechnology*, 2008, 10(2): 207–217
- Kochzius M, Seidel C, Antoniou A, *et al.* Identifying fishes through DNA barcodes and microarrays. *PLoS One*, 2010, 5(9): e12620
- Lara A, De Leon JLP, Rodriguez R, *et al.* DNA barcoding of Cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources*, 2010, 10(3): 421–430
- Le PQ. *China animal Scientific Osteichthyes*. Beijing: Science Press, 2000 [乐佩琦. 中国动物志, 硬骨鱼纲. 北京: 科学出版社, 2000]
- Liu CC, Gao X, Lin PC, *et al.* Fish community structure in Gezhouba reservoir. *Resources and Environment in the Yangtze Basin*, 2012, 21(7): 843–849 [刘春池, 高欣, 林鹏程, 等. 葛洲坝水库鱼类群落结构特征研究. 长江流域资源与环境, 2012, 21(7): 843–849]
- Liu SF, Li XR, Du TF, *et al.* DNA barcoding and electronic microarray for common fish species in Shandong coastal water. *Journal of Fishery Sciences of China*, 2016, 23(4): 777–790 [柳淑芳, 李献儒, 杜腾飞, 等. 山东近海习见鱼类 DNA 条形码及其电子芯片分析. 中国水产科学, 2016, 23(4): 777–790]
- Mabragana E, De Astarloa JMD, Hanner R, *et al.* DNA barcoding identifies Argentine fishes from marine and brackish waters. *PLoS One*, 2011, 6(12): e28655
- Ouyang XY, Mo BH, Yu HL, *et al.* DNA barcoding identification comparison on accuracy of DNA barcodes in mosquito identification. *Chinese Journal of Frontier Health and Quarantine*, 2007, 30(6): 349–352 [欧阳小艳, 莫帮辉, 余华丽, 等. DNA 条形码识别-DNA 条形码与 DNA 芯片识别蚊媒准确性的比较. 中国国境卫生检疫杂志, 2007, 30(6): 349–352]
- Peng JL, Wang XZ, Wang D, *et al.* Application of DNA barcoding based on the mitochondrial *CO I* gene sequences in classification of culter (Pisces: Cyprinidae). *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. 水生生物学报, 2009, 33(2): 271–276]
- Tamura K, Stecher G, Peterson D, *et al.* MEGA6: Molecular evolutionary genetics analysis Version 6.0. *Molecular Biology and Evolution*, 2013, 30(12): 2725–2729
- Yi Y, Yang Z, Zhang S. Ecological risk assessment of heavy

metals in sediment and human health risk assessment of heavy metals in fishes in the middle reaches of the Yangtze River basin. *Environmental Pollution*, 2011, 159(10):

2575–2585

Zhang JB, Hanner R. Molecular approach to the identification of fish in the South China Sea. *PLoS One*, 2012, 7(2): e30621

(编辑 冯小花)

DNA Barcoding and Electronic Microarrays for Common Fish Species in the Middle Reaches of the Yangtze River

GAO Xingchen¹, JIANG Wei^{1①}, LIU Shaoping²

(1. *Chinese Sturgeon Research Institute, China Three Gorges Corporation, Hubei Key Laboratory of Three Gorges Project for Conservation of Fishes, Yichang 443100*; 2. *Scientific Observing and Experimental Station of Fishery Resources and Environment in the Upper and Middle Reaches of Yangtze River, Ministry of Agriculture and Rural Affairs; Yangtze River Fishery Resources Institute, Chinese Academy of Fishery Sciences, Wuhan 430223*)

Abstract To develop a method for the rapid identification and classification of fishes in the middle reaches of the Yangtze River, 168 DNA sequences of the cytochrome oxidase subunit I (CO I) gene collected from GenBank. 64 common freshwater fish species in 7 orders, 11 families, and 50 genera were analyzed to test the efficacy of species identification using a DNA barcode microarray. The results showed that the intraspecific genetic distances ranged from 0 to 0.0776 (average, 0.0079). The interspecific distances ranged from 0.003 to 0.0190 (average, 0.084). The interspecific genetic distance was larger than the intraspecific distance, and all 168 sequences formed species units in a neighbor-joining phylogenetic tree, indicating that the DNA barcodes could be used to identify these 64 species. Based on the DNA microarray technology, 112 probes were selected from 43 species with high specificity, accounting for 67.2% of the total species. Thus, the DNA barcode microarray provided technical support and a new way to identify fish species in the middle reaches of Yangtze River basin.

Key words Middle reaches of Yangtze River; DNA barcode; Electronic microarray

① Corresponding author: JIANG Wei, E-mail: 106635291@qq.com