

DOI: 10.12264/JFSC2023-0153

## 基于线粒体 Cyt *b* 基因和 D-loop 区序列的长麦穗鱼遗传多样性研究

杨晓鸽<sup>1,2</sup>, 夏美琪<sup>1</sup>, 姜洁晨<sup>1</sup>, 汪健<sup>1</sup>, 陈娜<sup>1</sup>, 廖梦绮<sup>1</sup>, 连玉喜<sup>1,2</sup>

1. 安庆师范大学生命科学学院, 安徽 安庆 246133;
2. 水生生物保护与水生态修复安徽省高等学校工程技术研究中心, 安徽 安庆 246133

**摘要:** 为研究长麦穗鱼(*Pseudorasbora elongata*)的遗传多样性现状, 本研究利用线粒体 Cyt *b* 基因和 D-loop 区序列, 分析了长江中下游安徽境内皖南山区的闪里(SL)、历口(LK)和石台(ST) 3 个长麦穗鱼群体的遗传多样性及遗传结构。结果显示, 基于 Cyt *b* 和 D-loop 序列定义的单倍型数分别为 18 和 27, 整体单倍型多样性指数( $H_d$ )和核苷酸多样性指数( $\pi$ )分别为 0.792 和 0.01332、0.777 和 0.01140, 2 个标记均显示 ST 群体遗传多样性相对最低; 群体间的遗传距离为 0.00173~0.03615 (Cyt *b*)、0.00193~0.02639 (D-loop); 遗传分化指数( $F_{st}$ )和基因流( $N_m$ )均表明 ST 群体和 SL 及 LK 群体间存在显著的遗传分化; 分子方差分析(AMOVA)表明, 长麦穗鱼群体间的遗传变异(94.60%、90.69%)远高于群体内(5.40%、9.31%), 遗传变异主要来自群体间; 单倍型系统进化树及网络结构图分析显示, SL 和 LK 群体遗传关系较近, 聚为一支, 而 ST 群体单独聚为一支, 形成明显的地理谱系格局; 中性检验和核苷酸错配分布图表明长麦穗鱼整体近期历史上未发生扩张事件。总体来看, 长麦穗鱼遗传多样性丰富, 但各群体内的遗传多样性相对较为贫乏, 建议将 ST 群体作为一个保护管理单元进行优先保护, 本研究可为长麦穗鱼种质资源保护和可持续发展与利用提供参考。

**关键词:** 长麦穗鱼; Cyt *b* 基因; D-loop 区; 遗传多样性; 遗传结构

中图分类号: S931 文献标志码: A 文章编号: 1005-8737-(2023)08-1031-11

长麦穗鱼(*Pseudorasbora elongata*)隶属于鲤形目(Cypriniformes)、鲤科(Cyprinidae)、鮡亚科(Gobioninae)、麦穗鱼属, 是我国特有的小型鱼类。与其同属的麦穗鱼相比, 长麦穗鱼的分布范围极为狭窄, 仅分布于长江中下游水系的建德、石台和祁门, 以及西江水系的阳朔和桂林<sup>[1-4]</sup>。长麦穗鱼对生境要求严苛, 偏好清澈的缓流浅滩和卵砾石底质生境。受栖息地环境改变、过度捕捞等人类活动的影响, 长麦穗鱼野生资源量不断下降、分布区域逐渐萎缩等问题变得尤为突出, 已被《中国濒危动物红皮书: 鱼类》和《中国物种红色名录(第一卷)》列为“易危”种<sup>[5-6]</sup>。根据前期的调查结果及历史研究资料发现, 在长麦穗鱼的

历史分布区建德, 现已不见踪迹, 而漓江分布区在 20 世纪就已濒临绝迹<sup>[7-8]</sup>。目前, 长麦穗鱼仅在安徽境内长江流域的石台和祁门地区偶有发现。因此, 对安徽长江流域现存长麦穗鱼种质资源展开调查, 分析并厘清其遗传多样性和遗传结构现状, 评估其历史动态, 对长麦穗鱼种质资源的保护及后续的可持续发展极为重要。

遗传多样性是物种生存(适应)和发展(进化)的前提, 可反映出其进化历史及进化潜力。鱼类同其他脊椎动物一样, 其线粒体 DNA (mitochondrial DNA, mtDNA)呈闭合环状, 具有母系遗传、结构简单、进化速率快等特点, 是一种重要的分子标记<sup>[9]</sup>。mtDNA 中细胞色素 b (cytochrome *b*, Cyt *b*)是研究

收稿日期: 2023-06-20; 修订日期: 2023-07-13.

基金项目: 国家自然科学基金项目(31800357); 安徽省教育厅科研项目(2023AH030072)

作者简介: 杨晓鸽(1987-), 女, 博士, 副教授, 研究方向为濒危鱼类保护. E-mail: flyingdove2007@163.com

通信作者: 连玉喜, 副教授, 研究方向为鱼类生态学. E-mail: lianyx177@163.com

得最为清楚的蛋白质编码基因之一,其进化速率适中;而 mtDNA 控制区(control region displacement loop, D-loop, 又称 D-环区)不参与编码蛋白质,进化速率快,且与 Cyt *b* 基因在进化上存在差异<sup>[10-11]</sup>。因此,mtDNA Cyt *b* 基因和 D-loop 区常作为分子标记被广泛应用于鱼类的遗传多样性研究<sup>[12-15]</sup>。目前,已有学者对长麦穗鱼的 mtDNA 序列及结构做了描述,但其种群遗传结构及多样性未见系统报道<sup>[16]</sup>。研究发现,长麦穗鱼在形态学上与扁吻鮡属的鱼类有诸多相似特征,而在对麦穗鱼属系统发育关系的研究中发现,长麦穗鱼是最先分化出来的一支,与本属其他种亲缘关系相对较远<sup>[17]</sup>。据此,本研究拟采用长麦穗鱼 mtDNA Cyt *b* 基因及 D-loop 区序列对安徽皖南山区长麦穗鱼种群遗传

多样性现状和遗传结构进行分析和评估,以期为后续长麦穗鱼的系统进化、种质资源的鉴定及开发利用与可持续发展提供理论依据。

## 1 材料与方法

### 1.1 样品采集与 DNA 提取

长麦穗鱼样品于 2018 年 4 月至 2019 年 12 月用地笼网收集于长江中下游安徽境内皖南山区的闪里(SL)、历口(LK)和石台(ST) 3 地(图 1),共计 141 尾,取鳍条组织置于无水乙醇中保存。长麦穗鱼基因组 DNA 依据天根生化科技(北京)有限公司动物组织基因组 DNA 提取试剂盒(DP304-02)进行提取,基因组 DNA 经 1%的琼脂糖凝胶电泳检测后,保存于-20 °C 冰箱备用。

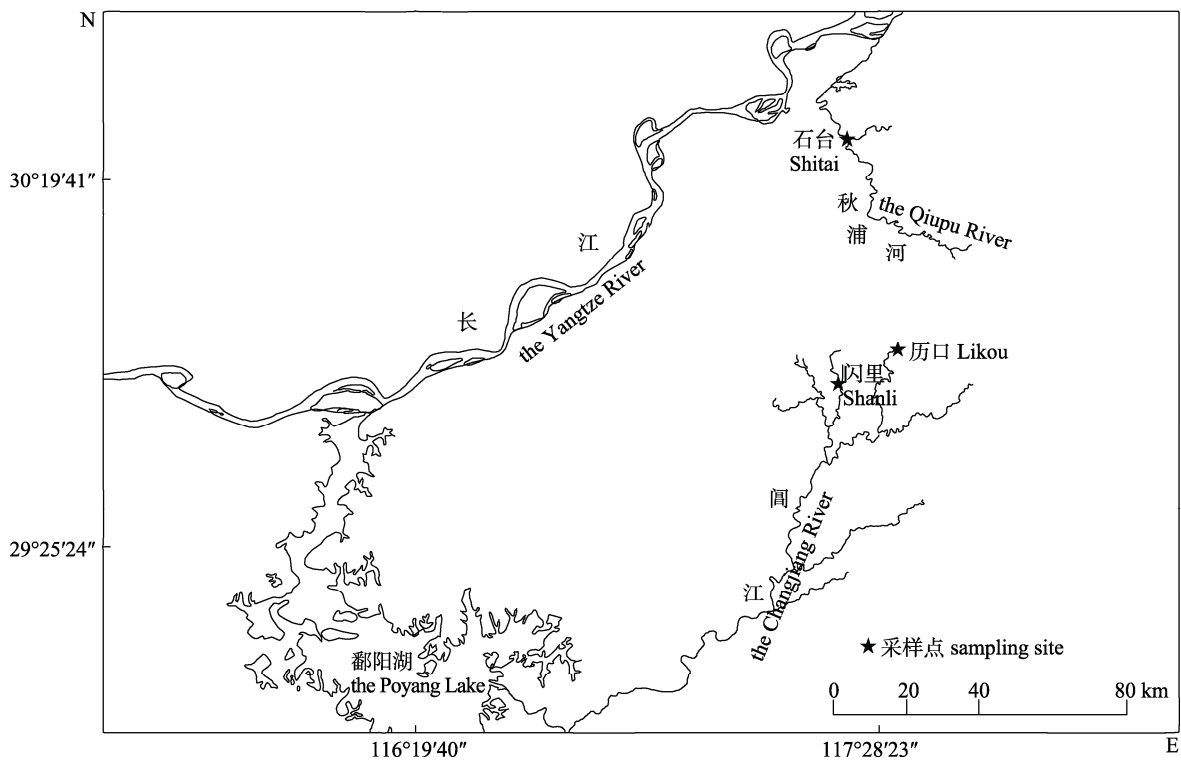


图 1 长麦穗鱼采样点分布图

Fig. 1 Sampling sites of *Pseudorasbora elongata*

### 1.2 PCR 扩增与测序

根据 NCBI 数据库中长麦穗鱼 mtDNA 全序列 (GenBank: KF051938), 利用 Primer Premier 5.0 设计引物用于 Cyt *b* 基因和 D-loop 区序列扩增。本研究所用扩增和测序的引物序列为, Cyt *b* F: 5'-ATGGCAAGCCTACGAAAACCC-3'、Cyt *b* R:

5'-AGGGCAAGCTCATTCTAGTGCTT-3'; D-loop F: 5'-TTAACTCCCACCCCTGGCTC-3'、D-loop R: 5'-CGGAGCTTTCTAGGGCCCAT-3'。PCR 反应体系 25  $\mu$ L, 包含 10 $\times$ PCR Buffer 2.5  $\mu$ L, MgCl<sub>2</sub> 2  $\mu$ L, dNTP 0.5  $\mu$ L, DNA 模板 1  $\mu$ L, 上、下游引物各 0.5  $\mu$ L, Taq DNA 聚合酶 0.5  $\mu$ L, ddH<sub>2</sub>O 补足体积。

扩增程序: 预变性 94 °C 3 min; 变性 94 °C 30 s, 52 °C (Cyt *b*) 或 54 °C (D-loop) 退火 30 s, 72 °C 延伸 1 min; 35 个循环, 72 °C 延伸 10 min。扩增产物经过 1.5% 琼脂糖凝胶电泳检测后, 送样测序。本研究引物合成及扩增产物测序均由北京擎科生物有限公司武汉分公司完成。为保证测序准确性, 所有样品均采用双向测序。

### 1.3 数据分析

所获得测序结果运用 Vector NTI 软件进行拼接, 序列整理后在 GenBank 中进行比对, 确认扩增片段即为本研究目的片段。利用 Clustal X2 软件对测序结果进行比对, 去除两端冗余序列。运用 DnaSP 5.0 软件获得基于 Cyt *b* 和 D-loop 序列的单倍型数(number of haplotypes,  $h$ )、单倍型多样性指数(haplotypes diversity,  $H_d$ )、核苷酸多样性指数(nucleotide diversity,  $\pi$ )、平均核苷酸差异指数(average number of nucleotide differences,  $K$ )以及多态位点数(number of polymorphic sites,  $S$ )<sup>[18]</sup>。使用 MEGA X 软件计算 Cyt *b* 基因和 D-loop 区序列的碱基组成、变异及保守位点、转换与颠换比值, 采用最大似然法(maximum likelihood, ML)构建系统发育进化树, 各分支置信度用 Bootstrap 重复检测 1000 次<sup>[19]</sup>。运用 Arlequin3.5 软件计算遗传分化指数( $F$ -statistics,  $F_{st}$ )和分子变异分析(analysis of molecular variance, AMOVA), 基因流( $N_m$ )由公式  $N_m \approx (1 - F_{st}) / 4F_{st}$  计算<sup>[20-21]</sup>。运用中介法(median joining, MJ)构建单倍型网络结构图(Network 5.0)。

## 2 结果与分析

### 2.1 长麦穗鱼线粒体 Cyt *b* 基因和 D-loop 区的序列特征

测序结果经过校正比对后, 分别获得长麦穗鱼 Cyt *b* 基因序列 1046~1049 bp 和 D-loop 序列 938~941 bp。通过 MEGA X 软件计算序列碱基组成, 结果显示皖南山区 3 个长麦穗鱼群体间 Cyt *b* 和 D-loop 序列 4 种碱基 A、T、G、C 的平均含量分别为 28.5%、32.3%、13.8%、25.4% 和 33.9%、34.0%、12.9%、19.2% (表 1), 两序列均出现明显 AT 偏移, A+T (60.8%、67.9%) 高于 G+C (39.2%、32.1%), 符合已知鱼类线粒体 DNA 序列特征。此

外, Cyt *b* 基因 4 种碱基在第 1、2 和 3 位密码子分布也出现偏倚性, 碱基 A 偏向第 1 位密码子 (34.7%), 碱基 T 偏向于第 3 位密码子 (38.8%)。Cyt *b* 基因序列有 3 个碱基插入或缺失, 包含 50 个变异位点, 其中单位点突变 11 个, 简约信息位点 39 个; 变异位点的转换/颠换比为 1.89, 转换以碱基 C/T 间为主。D-loop 区序列有 5 个碱基插入或缺失, 包含变异位点 43 个, 包括简约信息位点 37 个和单突变位点 6 个, 变异位点的转换和颠换比值为 1.43。

表 1 长麦穗鱼线粒体 Cyt *b* 基因和 D-loop 区序列的碱基组成

Tab. 1 Nucleotide composition of mtDNA Cyt *b* gene and D-loop region sequences of *Pseudorasbora elongata*

群体 population	Cyt <i>b</i>					D-Loop				
	A	T	G	C	A+T	A	T	G	C	A+T
闪里 SL	28.4	32.2	13.9	25.5	60.6	33.9	33.8	13.0	19.3	67.7
历口 LK	28.5	32.2	13.8	25.5	60.7	33.9	33.9	13.0	19.2	67.8
石台 ST	28.6	32.6	13.7	25.2	61.2	34.0	34.4	12.7	19.0	68.4
平均 Avg.	28.5	32.3	13.8	25.4	60.8	33.9	34.0	12.9	19.2	67.9

### 2.2 群体遗传多样性分析

使用 DNAsp5 对皖南山区 3 个地理群体长麦穗鱼的遗传多样性参数进行分析统计, 结果如表 2 所示。Cyt *b* 基因在长麦穗鱼 3 个群体的 137 个个体中共检测到 18 种单倍型, 3 个群体的单倍型多样性指数( $H_d$ )为 0.275~0.721, 长麦穗鱼整体  $H_d$  为 0.792; 3 个群体的核苷酸多样性指数( $\pi$ )为 0.00038~0.00172, 整体为 0.01332, 核苷酸多样性较高。核苷酸差异数( $K$ )在 3 个群体中的差异与核苷酸多样性一致。D-loop 区序列在 3 个群体 141 个个体中共检测到 27 种单倍型, 单倍型多样性指数为 0.053~0.773, 整体  $H_d$  为 0.777; 核苷酸多样性指数 0.00017~0.0026, 整体  $\pi$  为 0.01140; 3 个群体整体核苷酸差异数  $K$  为 10.623。其中, 长麦穗鱼 ST 群体 Cyt *b* 基因和 D-loop 区序列的  $H_d$  和  $\pi$  分别为 0.275 和 0.00038、0.053 和 0.00017, 均是 3 个群体中最低的。

### 2.3 群体间的遗传距离和遗传分化分析

基于 K2P 模型计算皖南山区长麦穗鱼群体内和群体间遗传距离(表 3)。长麦穗鱼 Cyt *b* 基因的

表 2 基于线粒体 Cyt *b* 基因和 D-loop 区序列的长麦穗鱼群体的遗传多样性参数和中性检验Tab. 2 Genetic diversity and neutral test of *Pseudorasbora elongata* based on the sequences of mtDNA Cyt *b* gene and D-loop region

群体 population	基因 gene	样本数 <i>n</i>	<i>S</i>	<i>h</i>	$H_d (\bar{x} \pm SD)$	$\pi (\bar{x} \pm SD)$	<i>K</i>	Tajima's <i>D</i>	Fu's $F_s$
闪里 SL	Cyt <i>b</i>	79	6	8	0.695±0.041	0.00138±0.00010	1.447	-2.68020	-0.95372
	D-loop	79	18	25	0.773±0.046	0.00269±0.00039	2.511	-0.91450	-6.10908**
历口 LK	Cyt <i>b</i>	24	12	8	0.721±0.084	0.00172±0.00049	1.797	-1.51212*	-2.13198
	D-loop	24	2	2	0.391±0.091	0.00083±0.00019	0.783	1.01428	2.47908
石台 ST	Cyt <i>b</i>	34	5	5	0.275±0.099	0.00038±0.00016	0.401	-1.77685**	-3.00615*
	D-loop	38	3	2	0.053±0.049	0.00017±0.00016	0.158	-1.49106*	-0.61060
整体 overall	Cyt <i>b</i>	137	50	18	0.792±0.023	0.01332±0.00116	13.920	1.82016	9.62978
	D-loop	141	43	27	0.777±0.026	0.01140±0.00080	10.623	1.40670	0.81160

注: 表中字母 *h* 表示单倍型数目; *S* 表示变异位点数目;  $H_d$  表示单倍型多样性指数;  $\pi$  表示核苷酸多样性指数; *K* 代表平均核苷酸差异数。\*和\*\*分别表示显著( $P<0.05$ )和极显( $P<0.01$ )显著偏离中性检验模型。

Note: *h*: the number of haplotypes; *S*: the number of variation sites;  $H_d$ : haplotype diversity;  $\pi$ : nucleotide diversity; *K*: the average of nucleotide difference. \*:  $P<0.05$ , \*\*:  $P<0.001$ .

表 3 基于线粒体 Cyt *b* 基因和 D-loop 区序列的长麦穗鱼群体遗传距离Tab. 3 Genetic distance within/among *Pseudorasbora elongata* populations based on the sequences of mtDNA Cyt *b* gene and D-loop region

群体 population	闪里 SL	历口 LK	石台 ST
闪里 SL	0.00138 (0.00270)		
历口 LK	0.00173 (0.00193)	0.00172 (0.00084)	
石台 ST	0.03391 (0.02639)	0.03615 (0.02519)	0.00038 (0.00011)

注: 对角线数值表示群体内遗传距离, 对角线下方表示群体间遗传距离, 括号内数值为基于 D-loop 的计算结果。

Note: the diagonal values show the genetic distance within population, the values of the genetic distance among populations are below the diagonal. The value of D-loop is in brackets.

分析结果显示, 3 个群体内的遗传距离为 0.00038~0.00172, 群体间的遗传距离为 0.00173~0.03615; 其中, ST 群体内的遗传距离最小(0.00038); ST 和 LK 间的遗传距离最大(0.03615); SL 和 LK 间的遗传距离最小(0.00173)。基于 D-loop 区序列的计算结果表明, ST 群体内的遗传距离最小(0.00011), SL 和 ST 群体间的遗传距离最大(0.02639)。

长麦穗鱼 3 个群体间遗传分化指数和基因流分析结果显示(表 4), 基于 Cyt *b* 基因和 D-loop 区序列长麦穗鱼群体间的固定指数  $F_{st}$  为 0.10688~0.97247、0.05332~0.98421, 并据此计算长麦穗鱼群体间的基因流  $N_m$  分别为 0.00707~2.08907、0.00401~4.43867, 表明 3 个长麦穗鱼群体间存在

表 4 基于线粒体 Cyt *b* 基因和 D-loop 区序列的长麦穗鱼群体间固定指数和基因流分析Tab. 4 Pairwise fixation indices and gene flow among *Pseudorasbora elongata* populations based on mtDNA Cyt *b* gene and D-loop region

群体 population	闪里 SL	历口 LK	石台 ST
闪里 SL		2.08907 (4.43867)	0.00851 (0.01904)
历口 LK	0.10688 (0.05332)		0.00707(0.00401)
石台 ST	0.96707 (0.92920)	0.97247 (0.98421)	

注: 对角线下方为遗传分化固定指数  $F_{st}$ ; 对角线上为基因流  $N_m$ 。括号内为基于 D-loop 的计算结果。

Note: Data below the diagonal shows the fixation index ( $F_{st}$ ) and gene flow ( $N_m$ ) is above the diagonal. The values in brackets are the results based on D-loop region.

不同程度的遗传分化, 且缺乏基因交流。相较而言, SL 和 LK 群体间分化程度较低, ST 和 SL 及 LK 群体间的遗传分化程度均较高。

长麦穗鱼群体间的分子方差分析(AMOVA)结果如表 5 所示, 基于 Cyt *b* 基因和 D-loop 区序列, 长麦穗鱼群体间的遗传变异分别占 94.60%、90.69%, 群体内的遗传变异占 5.40%、9.31%, 显示出长麦穗鱼群体间的遗传变异远高于群体内, 表明长麦穗鱼的遗传变异主要来自于群体间。

## 2.4 单倍型分布与系统进化分析

经 ClustalX 软件比对, 基于 Cyt *b* 基因和 D-loop 区序列运用 DnaSP 软件分别定义了 18 和 27 种单倍型。在 Cyt *b* 基因的单倍型中, Hap2 是优势单倍型, 占比高达 37.2%, 为 SL 和 LK 群体

表 5 长麦穗鱼群体遗传差异的分子方差分析(AMOVA)  
Tab. 5 Analysis of molecular variance (AMOVA) of *Pseudorasbora elongata*

变异来源 source of variation	自由度 <i>df</i>	平方和 sum of squares	变异组成 variance component	变异百分比/% percentage of variation
群体间 among populations	2 (2)	862.869 (634.751)	10.93386 (7.68309) Va	94.60 (90.69)
群体内 within populations	134 (138)	83.715 (108.859)	0.62474 (0.78883) Vb	5.40 (9.31)
总计 total	136 (140)	946.584 (743.610)	11.55859 (8.47192)	

注: 括号内的数值为基于 D-loop 区序列的计算结果; Va 和 Vb 分别表示群体间和群体内变异。

Note: The values in brackets are the results based on the D-loop region. Va and Vb represent variance components among and within populations respectively.

所共有; Hap1、Hap4、Hap7、Hap10、Hap11 是 SL 群体特有的单倍型; Hap6、Hap8、Hap9、Hap12、Hap13 为 LK 群体所独享; 值得注意的是 ST 群体的 5 种单倍型均为该群体所特有, 且 Hap14 是 ST 群体的优势单倍型(表 6)。在 D-loop 区序列定义

的单倍型中, Hap1 是优势单倍型, 为 SL 和 LK 群体共有; 值得注意的是 SL 群体独享的单倍型种数高达 23 种, 而 ST 群体仅有 Hap26 和 Hap27 两种单倍型, 均为该群体所特有, 且 Hap26 是 ST 群体的优势单倍型(表 6)。

表 6 基于线粒体 Cyt *b* 基因和 D-loop 区序列的长麦穗鱼 3 个群体的单倍型分布  
Tab. 6 Haplotype distribution in three *Pseudorasbora elongata* populations based on mtDNA Cyt *b* gene and D-loop region

基因 gene	群体 population	样本量 sampling size	单倍型数 haplotype number	单倍型分布 haplotype distribution
Cyt <i>b</i>	闪里 SL	79	8	Hap1(15), Hap2(39), Hap3(8), Hap4(12), Hap5(2), Hap7(1), Hap10(1), Hap11(1)
	历口 LK	24	8	Hap2(12), Hap3(5), Hap5(2), Hap6(1), Hap8(1), Hap9(1), Hap12(1), Hap13(1)
	石台 ST	34	5	Hap14(29), Hap15(1), Hap16(1), Hap17(1), Hap18(1)
D-loop	闪里 SL	79	25	Hap1(35), Hap2(2), Hap3(1), Hap4(1), Hap5(1), Hap6(3), Hap7(1), Hap8(3), Hap9(1), Hap10(4), Hap11(11), Hap12(1), Hap13(1), Hap14(1), Hap15(1), Hap16(1), Hap17(1), Hap18(1), Hap19(2), Hap20(1), Hap21(1), Hap22(1), Hap23(1), Hap24(1), Hap25(1)
	历口 LK	24	2	Hap1(19), Hap6(5)
	石台 ST	38	2	Hap26(37), Hap27(1)

以同属的麦穗鱼(*Pseudorasbora parva*, 基因登录号: JF802126.1)及扁吻鲷属的扁吻鲷(*Pungtungia herzi*, 基因登录号: AB239598.1)为外群, 基于长麦穗鱼 Cyt *b* 基因定义的 18 种单倍型和 D-loop 区序列定义的 27 种单倍型, 分别构建 ML 系统进化树(图 2a、2b)。由系统进化树可知, 基于 Cyt *b* 基因(图 2a)和 D-loop 区(图 2b)序列构建的两个长麦穗鱼系统发育树均可明显地分为 2 个分支: 一个分支由 ST 群体的单倍型构成; 另一个分支由 SL 和 LK 群体的单倍型构成。这与上述 ST 和 SL、LK 群体间的遗传距离相对较远, 群体间的遗传分化指数较高相符。另外, 值得注意的是, 系统进化显示出长麦穗鱼与扁吻鲷聚在一起, 而不是同属的麦穗鱼, 表明其与扁吻鲷的亲缘

关系更近。

基于中介法(medium-join, MJ), 运用 Network5.0 软件构建长麦穗鱼 Cyt *b* 基因和 D-loop 区单倍型网络结构图。基于 Cyt *b* 基因 18 种单倍型构建的网格图如图 3a 所示, Hap2 和 Hap3 单倍型较为古老, 位于网格图中心位置, 因 3 个节点而衍生出来的单倍型呈辐散状分布于其周围; ST 群体的优势单倍型 Hap14 由 SL 群体 Hap1 衍生而来。基于 D-loop 序列 27 种单倍型构建的网格图如图 3b, 网格图整体以 Hap1 单倍型为中心, 各群体的单倍型都直接或间接由 Hap1 突变后形成, 呈辐散状分布于其周围。总体来看, 长麦穗鱼单倍型网络图呈现的地理分布格局与系统进化树结果类似。

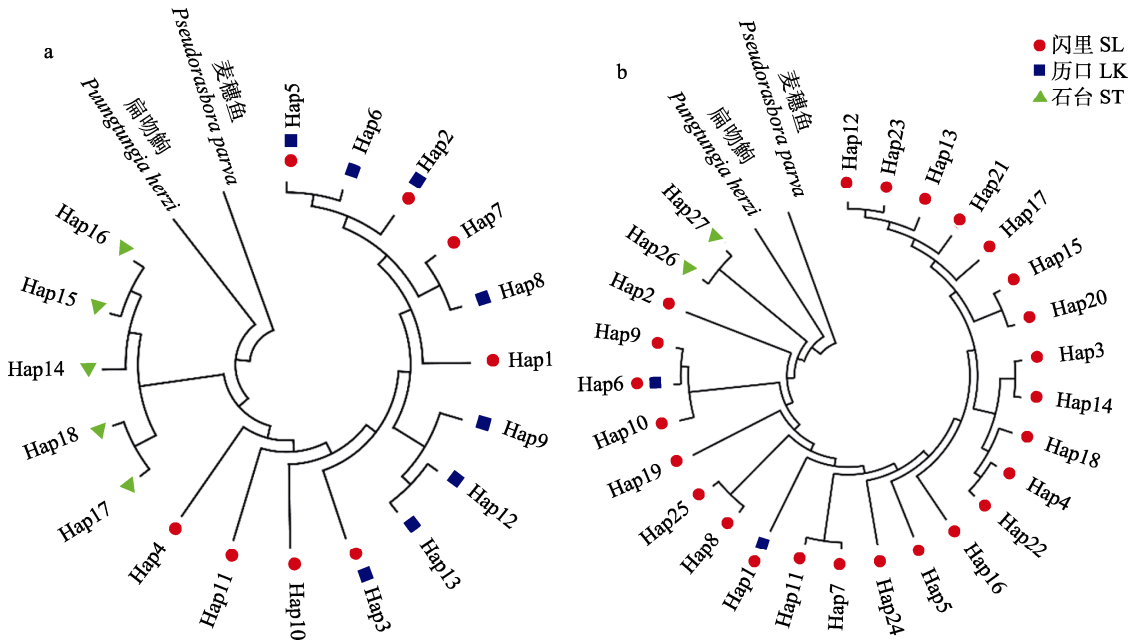


图 2 基于线粒体 Cyt *b* 基因(a)和 D-loop 区(b)序列的长麦穗鱼单倍型系统进化树  
 Fig. 2 The phylogenetic tree based on haplotypes of Cyt *b* gene (a) and D-loop region (b) in *Pseudorasbora elongata*

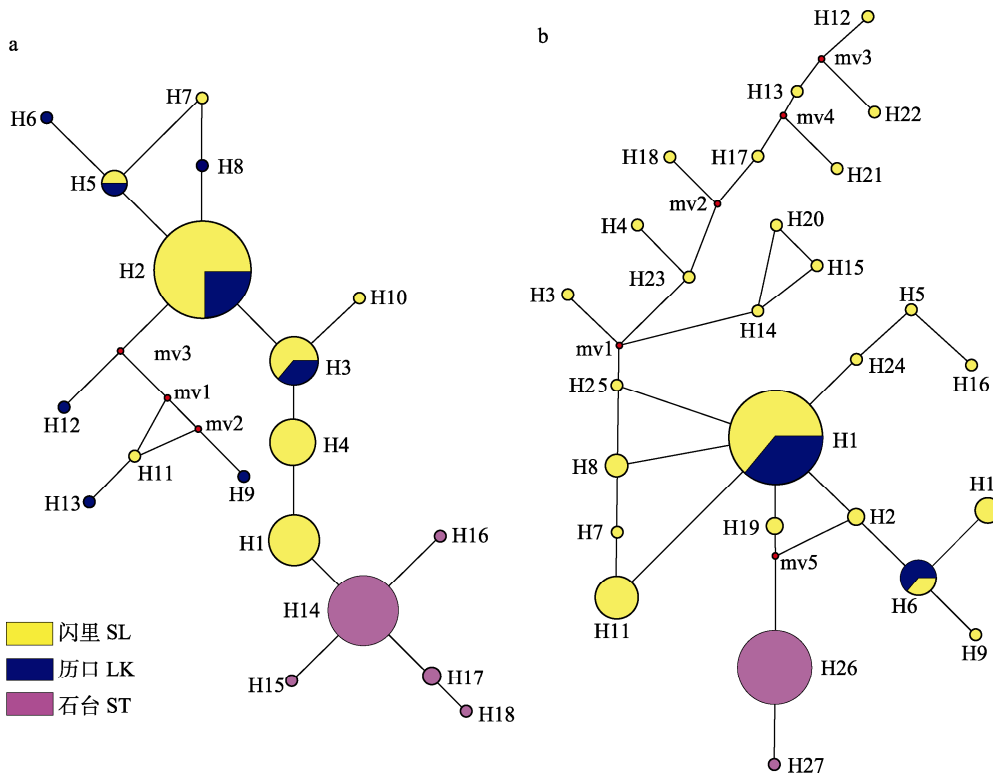


图 3 基于长麦穗鱼线粒体 Cyt *b* 基因(a)和 D-loop 区(b)单倍型的网络结构图  
 Fig. 3 The haplotype network based on Cyt *b* gene (a) and D-loop region (b) in *Pseudorasbora elongata*

2.5 种群历史动态

将皖南山区长麦穗鱼 3 个群体作为整体基于 Cyt *b* 基因和 D-loop 区序列进行中性检验, 结果显示 Tajima'*D* 值和 Fu's *F<sub>s</sub>* 值均为正值(表 2), 且

核苷酸错配分布分布图为多峰(图 4d、4h), 表明皖南山区长麦穗鱼符合中性进化假设, 历史上未发生过种群扩张。基于 Cyt *b* 基因 3 个群体的 Tajima'*D* 值和 Fu's *F<sub>s</sub>* 值均为负值, 且对于 ST 群

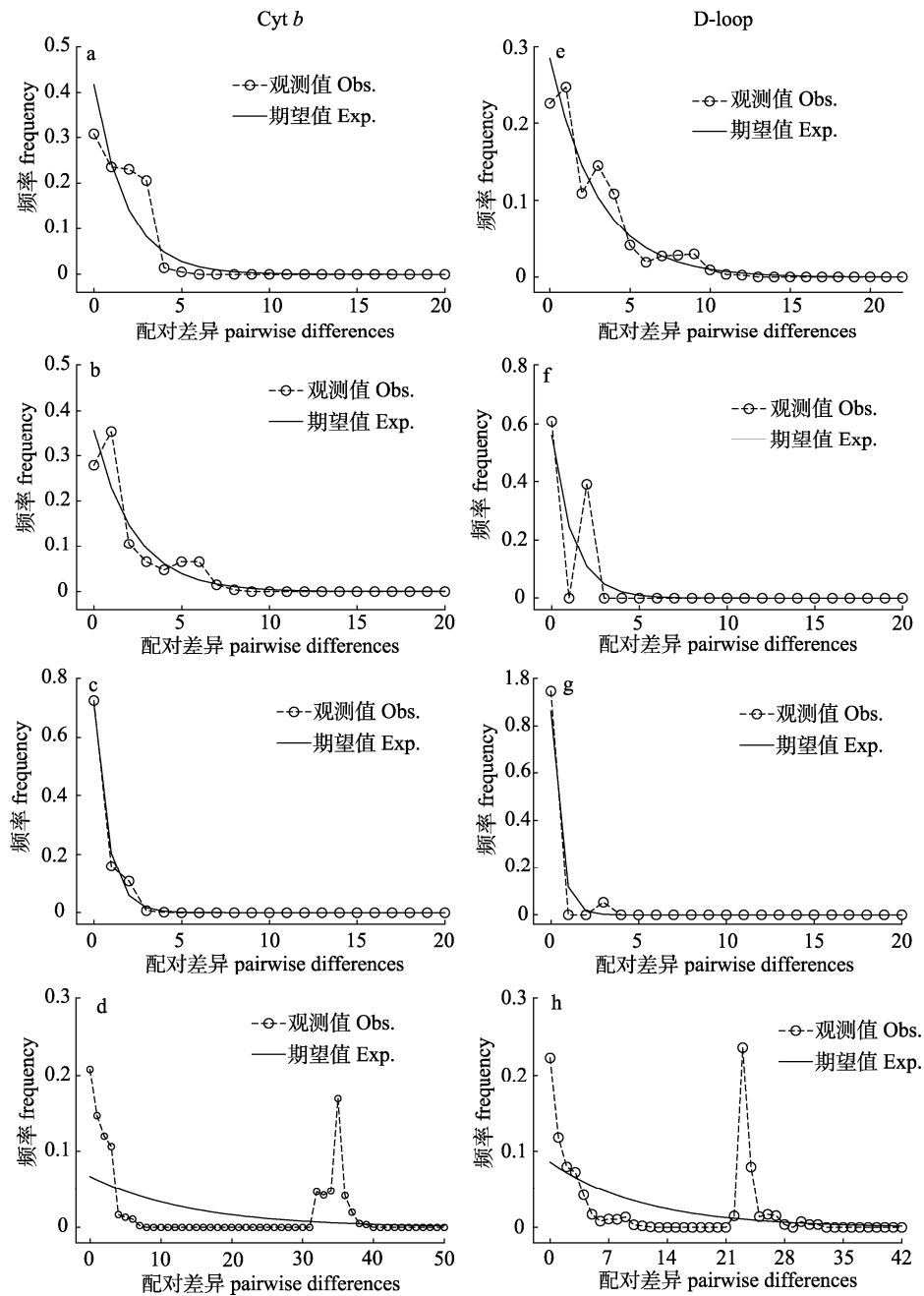


图 4 基于线粒体 Cyt b 基因和 D-loop 区序列的长麦穗鱼核苷酸错配分布图

a. 闪里, b. 历口, c. 石台, d. 整体(Cyt b); e. 闪里, f. 历口, g. 石台, h. 整体(D-loop).

Fig. 4 Mismatch distribution of *Pseudorasbora elongata* based on mtDNA Cyt b gene and D-loop region sequences  
a. SL, b. LK, c. ST, d. Overall (Cyt b); e. SL, f. LK, g. ST, h. Overall (D-loop).

体来说 Tajima's  $D$  值( $P < 0.01$ )和 Fu's  $F_s$  值( $P < 0.05$ )统计学分析存在差异, 偏离中性检验模型, 其核苷酸错配分布图呈现单峰(图 4c), 表明 ST 群体在历史上发生过种群扩张; SL 和 LK 群体核苷酸不配对图呈现多峰(图 4a、4b), 说明其未发生扩张。基于 D-loop 区序列, ST 群体 Tajima's  $D$  值和 Fu's  $F_s$  值均为负值, 且 Tajima's  $D$  值统计分析中存在显著

差异( $P < 0.05$ ), 其核苷酸不配对图呈单峰, 也说明其历史上发生过扩张(图 4g)。而 LK 群体 Tajima's  $D$  值和 Fu's  $F_s$  值为正值, SL 群体 Tajima's  $D$  值和 Fu's  $F_s$  值为负值, 但 Tajima's  $D$  值不存在显著性差异( $P > 0.05$ ), 二者核苷酸错配分布图均呈现多峰(图 4e、4f), 说明历史上未发生群体扩张事件。

### 3 讨论

研究种群的遗传多样性可揭示其起源与进化历史,亦可反映出其进化潜力。遗传多样性研究是评价种群资源状况以及开展保护工作的重要基础与依据,高水平的遗传多样反映出种群对环境有较强的适应能力,受环境压力影响小,该种群越容易延续;相反,遗传多样性水平过低则可能导致种群遭遇瓶颈甚至灭亡<sup>[22-23]</sup>。

#### 3.1 长麦穗鱼种群遗传多样性

本研究以 mtDNA 的 Cyt *b* 基因和 D-loop 区为标记对 3 个长麦穗鱼群体遗传多样性现状展开了调查和研究。通过 PCR 获得 SL、LK、ST 3 个长麦穗鱼群体线粒体 Cyt *b* 基因和 D-loop 区同源片段序列,分析显示序列碱基含量 A+T 均高于 G+C,出现明显的 AT 偏倚与反 G 的偏歧现象,这与已知包括鱼类在内的动物线粒体基因组核苷酸分布不均一情况相一致<sup>[12-15,23]</sup>。D-loop 区序列不编码蛋白质,且在 mtDNA 中进化速率最高,变异通过母系遗传而保留,以此为遗传标记获得的单倍型数往往更多。在本研究中,D-loop 区序列定义了 27 个单倍型,高于 Cyt *b* 基因标记定义的单倍型数(18 个)。单倍型多样性指数( $H_d$ )和核苷酸多样性指数( $\pi$ )是常用来评估物种遗传多样性高低的指标。根据 Grant<sup>[24]</sup>和 Bowen 等<sup>[25]</sup>提出来的分类标准( $H_d=0.5$ 、 $\pi=0.005$ ),将遗传多样性划分为 4 种模式。本研究中,长麦穗鱼整体的  $H_d$  为 0.792 (Cyt *b*)、0.777 (D-loop), $\pi$  为 0.01332 (Cyt *b*)、0.01140 (D-loop)符合上述高单倍型多样性和低核苷酸多样性的模式。由于核苷酸多样性的积累比单倍型更漫长,推测符合该遗传多样性模式的种群可能经历了瓶颈效应或建群效应后扩张,同时又无充足时间去积累核苷酸变异<sup>[24-25]</sup>。ST 群体长麦穗鱼  $H_d$  为 0.275 (Cyt *b*)、0.053 (D-loop), $\pi$  为 0.00038 (Cyt *b*)、0.00017 (D-loop),均远低于上述划分标准。这一结果表明 ST 群体遗传多样性处于较低水平,该群体极有可能遭受过瓶颈效应,亦或是基础群体较小,遗传多样性仅有单一或极少数群体形成,这也反映出 ST 群体资源量严重枯竭的现状。此外,本研究发现基于 D-loop 区序列获得的遗传变异数及遗传多样性指数均低于

Cyt *b* 基因的结果。一般情况下,D-loop 区进化速率快,获得的种群的遗传变异通常大于 Cyt *b* 基因,但赵亮等<sup>[26]</sup>的研究种也发现的 mtDNA 控制区的变异速率低于 Cyt *b* 基因。综上,对于长麦穗鱼遗传多样性现状应加强其保护力度,着力增加其资源量,以提高其遗传多样性水平,避免长麦穗鱼种质资源进一步衰退。

#### 3.2 长麦穗鱼种群的遗传结构与遗传分化

不同物种间或同种不同种群间的遗传差异程度通常用遗传距离来衡量。本研究中 3 个长麦穗鱼群体间的遗传距离为 0.00173~0.03615 (Cyt *b*)、0.00193~0.02639 (D-loop),相较而言,SL 和 LK 群体间的遗传距离较小,ST 群体与 SL 及 LK 间的遗传距离均较大,这与其地理分布格局相一致。基于 Shaklee 等<sup>[27]</sup>提出的鱼类在属(0.9)、种(0.3)、种群(0.05)三级水平上的划分依据,本研究 3 个长麦穗鱼群体未达到种群划分标准(群体间的遗传距离均小于 0.05),说明 3 个群体可能由同一祖先种群扩散形成。鱼类因地理隔离或在同一水域中栖息环境不同而限制种群的基因交流,这是种群分化的重要原因,并由此呈现出与其分布水系相吻合的遗传分化格局<sup>[28]</sup>。遗传分化指数  $F_{st}$  是评估遗传分化程度的重要指标,在实际研究中,根据 Wright 的分类标准:当  $0 < F_{st} < 0.05$  时,表明群体间遗传差异很小;  $0.05 < F_{st} < 0.15$ ,表明群体间遗传分化达到中等程度水平;  $0.15 < F_{st} < 0.25$ ,表明存在较大的遗传分化;  $F_{st} > 0.25$ ,则遗传分化很大<sup>[28-29]</sup>。本研究中 SL 和 LK 群体间  $F_{st}$  为 0.10688 (Cyt *b*)、0.05332 (D-loop),处于中等程度的遗传分化;而 ST 和 LK 及 ST 和 LK 群体间的  $F_{st}$  值均大于 0.9,表明该群体间呈现高度遗传分化格局。此外,基因流也是评估种群遗传结构和遗传分化的重要指标,本研究中 SL 和 LK 群体间的基因流  $N_m$  值大于 1;而 ST 和 SL 及 ST 和 LK 群体间的  $N_m$  值均远远小于 1。研究认为, $N_m < 1$  则表明群体间遗传分化较大; $N_m > 1$  表明群体间地理距离很近或有某种渠道可以进行基因交流,群体间的遗传分化较小; $N_m > 4$  则说明群体间个体可以随机交配<sup>[30-31]</sup>。分子方差分析 AMOVA 结果显示长麦穗鱼群体间的遗传变异高达 90%以上,说明群体间的变异是



长麦穗鱼遗传变异的主要来源。综上, 造成 3 个长麦穗鱼群体呈现上述遗传分化格局的原因与其分布的水系密切相关, SL 和 LK 群体同属于闾江上游支流, 有基因交流的机会; 而 ST 群体分布于秋浦河, 由于水系的隔离限制了 ST 群体与其他两个群体间的基因交流, 从而与其他群体间发生了明显的遗传分化。

基于 Cyt *b* 基因和 D-loop 区序列定义的单倍型构建了 ML 系统进化树。从系统进化角度来看, 3 个长麦穗鱼群体形成了两个明显的地理聚群: 一支由 SL 和 LK 群体单倍型组成且这两个群体单倍型互相混杂未形成明显的系统进化关系; 另一支由 ST 群体的单倍型构成且与其他 2 个群体不共享, 表明 ST 群体与 SL 和 LK 群体缺乏基因交流, 形成了明显的地理遗传结构, 这与上述遗传分化格局结果相一致。此外, 单倍型 Hap2 (Cyt *b*)、Hap1 (D-loop) 均为 SL 和 LK 群体的优势单倍型且在网络结构图中处于中心位置, 网络结构图呈现出的单倍型地理分布格局与 ML 系统进化树结果一致。本研究中, 值得注意的是本研究系统进化表明长麦穗鱼和同属的麦穗鱼亲缘关系较远, 反而与扁吻鲃属的扁吻鲃亲缘关系更近(图 2)。长麦穗鱼在形态上与扁吻鲃有相似特征, 如二者体型均近似圆筒状, 而同属的麦穗鱼体型侧扁; 长麦穗鱼和扁吻鲃体侧自吻端沿侧线至尾鳍有黑色纵纹, 而麦穗鱼则无此特征<sup>[1]</sup>。因此, 建议后续对长麦穗鱼的系统进化可做进一步深入研究。

### 3.3 长麦穗鱼种群的历史动态与保护

利用核苷酸错配分布图与中性检验推测种群是否经历过扩张事件。若 Tajima'*D* 和 Fu's *F<sub>s</sub>* 呈现负值, 且具备统计学意义, 则表明该种群显著偏离中性检验模型, 反映出种群历史上可能经历过扩张事件<sup>[32]</sup>。此外, 若核苷酸错配分布图呈现泊松分布的单峰, 则表明种群可能经历了种群扩张事件<sup>[33]</sup>。本研究中核苷酸错配分布图及中性检验结果表明长麦穗鱼整体历史上近期未发生过明显的群体扩张。但 ST 群体的核苷酸错配分布图呈现单峰, Tajima'*D* 和 Fu's *F<sub>s</sub>* 检验存在显著性差异, 偏离中性进化假设, 说明 ST 群体历史上可能经历过种群扩张事件或者近期遭受了瓶颈效应。

研究鱼类的遗传多样性现状、遗传结构与谱系地理格局可为鱼类保护策略及渔业管理措施的制定提供参考。长麦穗鱼是我国特有的小型鱼类, 近年来受环境变化及人为因素的影响, 其野生资源量不断下降。本研究发现长江中下游安徽境内长麦穗鱼遗传多样性总体较为丰富, 但各群体内的遗传多样性相对较为贫乏, 3 个群体呈现出 2 个主要的谱系关系。其中, ST 群体的遗传多样性最低, 且其遗传距离、遗传分化指数等均与其他两个群体存在一定程度的隔离。根据进化显著单元 (evolutionarily significant unit, ESU) 设置的原则<sup>[34]</sup>, 建议将长麦穗鱼 ST 群体作为一个保护管理单元进行优先保护, SL 和 LK 群体作为一个整体进行管理和保护, 加强对长麦穗鱼种质资源遗传多样性的监测, 长期追踪其变化, 适时地制定科学合理的保护和利用措施以保证其资源的可持续性发展。

### 参考文献:

- [1] Chen Y Y. Zoology of China: Osteichthys Cypriniformes (Middle Volume)[M]. Beijing: Science Press, 1998: 265. [陈宜瑜. 中国动物志硬骨鱼纲鲤形目(中卷)[M]. 北京: 科学出版社, 1998.]
- [2] Fisheries Research Institute of Guangxi Zhuang Autonomous Region, Institute of Zoology, Chinese Academy of Science. Freshwater Fishes of Guangxi[M]. Guilin: Guangxi Nation Publishing House, 1981: 118. [广西壮族自治区水产研究所, 中国科学院动物研究所. 广西淡水鱼类志[M]. 桂林: 广西人民出版社, 1981: 118.]
- [3] Wang W J. Spatial pattern in the structure and diversity of the taxonomic and functional organizations of stream fish in the mountains of southern Anhui[D]. Wuhu: Anhui Normal University. [王文剑. 皖南山区溪流鱼类分类群与功能群结构及其多样性的空间格局[D]. 芜湖: 安徽师范大学.]
- [4] Cai D S, Zhao X G, Zhu Y, et al. Investigation on fish resources and analysis of species diversity in Lijiang River[J]. Journal of Guangxi Normal University (Natural Science Edition), 2009, 27(2): 130-136. [蔡德所, 赵湘桂, 朱瑜, 等. 漓江鱼类资源调查及物种多样性分析[J]. 广西师范大学学报(自然科学版), 2009, 27(2): 130-136.]
- [5] Yue P Q, Chen Y Y. China Red Data Book of Endangered Animals: Pisces[M]. Beijing: Science Press, 1998: 247. [乐佩琦, 陈宜瑜. 中国濒危动物红皮书: 鱼类[M]. 北京: 科学出版社, 1998: 247.]
- [6] Wang S, Xie Y. China Species Red List[M]. Beijing: Higher Education Press, 2004. [汪松, 解焱. 中国物种红色名录-第

- 一卷-红色名录:中英文本[M].北京:高等教育出版社,2004.]
- [7] Kong D P, Cui G H, Yang J X. Threatened fishes of the world: *Pseudorasbora elongata* Wu, 1939 (Cyprinidae)[J]. *Environmental Biology of Fishes*, 2006, 76(1): 69-70.
- [8] Wu X W. Fauna of Cyprinidae in China [M]. Shanghai: Shanghai Publisher of Science and Technology, 1982: 462-464. [伍献文. 中国鲤科鱼类志[M]. 上海: 上海科学技术出版社, 1982: 462-464.]
- [9] Guo X H, Liu S J, Liu Q, et al. New progresses on mitochondrial DNA in fish[J]. *Acta Genetica Sinica*, 2004, 31(9): 983-1000. [郭新红, 刘少军, 刘巧, 等. 鱼类线粒体 DNA 研究新进展[J]. *遗传学报*, 2004, 31(9): 983-1000.]
- [10] Gray M W, Burger G, Lang B F. Mitochondrial evolution[J]. *Science*, 1999, 283(5407): 1476-1481.
- [11] Yu X R, Qiu X M, Liu X Y, et al. Application of research on marine animals' population genetic structure using mitochondrial DNA polymorphism[J]. *Biotechnology Bulletin*, 2011(10): 49-54. [于旭蓉, 仇雪梅, 柳晓瑜, 等. 线粒体 DNA 多态性在海洋动物群体遗传结构研究中的应用[J]. *生物技术通报*, 2011(10): 49-54.]
- [12] Cheng Q Q, Lyu H, Pang J H, et al. Genetic diversity of four wild *Siniperca kneri* populations in the Yangtze River[J]. *Journal of Fishery Sciences of China*, 2019, 26(4): 774-782. [程起群, 吕浩, 逢娇慧, 等. 长江流域 4 个野生大眼鲈群体的遗传多样性分析[J]. *中国水产科学*, 2019, 26(4): 774-782.]
- [13] Chen L, He P P, Wei P Y, et al. Genetic structure of three populations of rice flower carp (*Cyprinus carpio*) in Guangxi Zhuang Autonomous Region based on mitochondrial D-loop region and Cyt *b* gene[J]. *Acta Hydrobiologica Sinica*, 2021, 45(1): 54-59. [程磊, 何莘萍, 韦媛媛, 等. 基于线粒体 D-loop 区和 Cyt *b* 基因分析广西禾花鲤三个群体遗传结构[J]. *水生生物学报*, 2021, 45(1): 54-59.]
- [14] Zhao W H, Yi S K, Su J X, et al. Genetic diversity of *Triplophysa yarkandensis* populations in Tarim River Basin in Xinjiang[J]. *Acta Hydrobiologica Sinica*, 2022, 46(3): 364-374. [赵文浩, 易少奎, 苏君晓, 等. 新疆塔里木河叶尔羌高原鳅遗传多样性研究[J]. *水生生物学报*, 2022, 46(3): 364-374.]
- [15] Yu K, An M, Huang S, et al. Analysis of genetic diversity of little scale mandarin fish *Coreoperca whiteheadi* population based on cytb and D-loop genes[J]. *Fisheries Science*, 2020, 39(6): 894-901. [余科, 安苗, 黄胜, 等. 基于 Cytb 和 D-Loop 基因对中国少鳞鳊种群遗传多样性的分析[J]. *水产科学*, 2020, 39(6): 894-901.]
- [16] He L P, Yang S B, Zheng D H, et al. Complete mitochondrial genome of *Pseudorasbora elongata* (Cypriniformes: Cyprinidae): Sequencing and analysis[J]. *Mitochondrial DNA*, 2014, 25(6): 433-434.
- [17] Ma J. Study on phylogenetic relationship and genetic diversity of Triticeae population[D]. Guangzhou: South China Normal University, 2014. [马晶. 麦穗鱼属系统发育关系及麦穗鱼种群遗传多样性的研究[D]. 广州: 华南师范大学, 2014.]
- [18] Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data[J]. *Bioinformatics*, 2009, 25(11): 1451-1452.
- [19] Kumar S, Stecher G, Li M, et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms[J]. *Molecular Biology and Evolution*, 2018, 35(6): 1547-1549.
- [20] Excoffier L, Lischer H E L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows[J]. *Molecular Ecology Resources*, 2010, 10(3): 564-567.
- [21] Wright S. The genetical structure of populations[J]. *Annals of Eugenics*, 1951, 15(4): 323-354.
- [22] Ward R D. Genetics in fisheries management[M]//Marine Genetics. Dordrecht: Springer Netherlands, 2000: 191-201.
- [23] Chen K, Wang W F, Chen X L, et al. Analysis of genetic diversity among five populations of *Luttraria maxima* based on mtDNA Cytb and D-Loop sequences[J]. *Marine Fisheries*, 2020, 42(2): 149-160. [陈康, 王伟峰, 陈秀荔, 等. 基于线粒体 Cytb 和 D-Loop 序列的大獭蛤 5 个地理群体遗传多样性研究[J]. *海洋渔业*, 2020, 42(2): 149-160.]
- [24] Grant W, Bowen B W. Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation[J]. *Journal of Heredity*, 1998, 89(5): 415-426.
- [25] Bowen W, Bass A L, Rocha L A, et al. Phylogeography of the trumpettefishes (Aulostomus): Ring species complex on a global scale[J]. *Evolution; International Journal of Organic Evolution*, 2001, 55(5): 1029-1039.
- [26] Zhao L, Xie B G, Liu Z J, et al. Molecular structure and DNA substitution rate of the mitochondrial control region and cytochrome *b* in Taihu salangid, *Neosalanx taihuensis*[J]. *Chinese Journal of Zoology*, 2010, 45(2): 27-38. [赵亮, 谢本贵, 刘志瑾, 等. 太湖新银鱼线粒体 D-loop 和 Cyt *b* 片段序列结构与进化速率比较[J]. *动物学杂志*, 2010, 45(2): 27-38.]
- [27] Shaklee J B, Tamaru C S, Waples R S. Speciation and evolution of marine fishes studied by the electrophoretic analysis of proteins[J]. *Pacific Science*, 1982, 36(2): 141-157.
- [28] Wright S. Evolution in Mendelian populations[J]. *Genetics*, 1931, 16(2): 97-159.
- [29] Wright S. Evolution and the Genetics of Populations: Variability within and among Natural Population. Chicago: University of Chicago Press, 1978: 79-103.
- [30] Qu R Z, Hou L, Lu H L, et al. The gene flow of population genetic structure[J]. *Hereditas(Beijing)*, 2004, 26(3): 377-382. [曲若竹, 侯林, 吕红丽, 等. 群体遗传结构中的基因流[J]. *遗传*, 2004, 26(3): 377-382.]
- [31] Petit R J, Excoffier L. Gene flow and species delimitation[J]. *Trends in Ecology & Evolution*, 2009, 24(7): 386-393.

- [32] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism[J]. *Genetics*, 1989, 123(3): 585-595.
- [33] Márcia Barbosa A, Real R, Muñoz A R, et al. New measures for assessing model equilibrium and prediction mismatch in species distribution models[J]. *Diversity and Distributions*, 2013, 19(10): 1333-1338.
- [34] Moritz C. Defining 'evolutionarily significant units' for conservation[J]. *Trends in Ecology & Evolution*, 1994, 9(10): 373-375.

## Genetic diversity of *Pseudorasbora elongata* based on mitochondrial Cyt *b* gene and D-loop region sequences

YANG Xiaoge<sup>1,2</sup>, XIA Meiqi<sup>1</sup>, JIANG Jiechen<sup>1</sup>, WANG Jian<sup>1</sup>, CHEN Na<sup>1</sup>, LIAO Mengqi<sup>1</sup>, LIAN Yuxi<sup>1,2</sup>

1. College of Life Science, Anqing Normal University, Anqing 246133, China;

2. Engineering Technology Research Center for Aquatic Organism Conservation and Water Ecosystem Restoration in University of Anhui Province, Anqing 246133, China

**Abstract:** The slender top-mouth gudgeon *Pseudorasbora elongata* is a rare fish belonging to the family Cyprinidae. This species is endemic to China and distributed in restricted regions around the tributaries of the middle and lower Yangtze and Xijiang Rivers. *P. elongata* has suffered an increasing decline and has been vulnerable to extinction in the wild due to illegal and overfishing, water project construction, and other anthropogenic factors that destroyed their natural habitat and spawning grounds over the last few decades. Therefore, it is vital to assess the genetic diversity and structure of *P. elongata* for conserving and utilizing this fish effectively. In this study, we analyzed three *P. elongata* populations from Shanli (SL), Likou (LK), and Shitai (ST) in the Anhui section along the Yangtze River based on mitochondrial Cyt *b* gene and D-loop region sequences. The results showed that the total haplotypes defined by Cyt *b* gene and D-loop region in all populations were 18 and 27, respectively; the corresponding overall haplotype diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ) were 0.792 and 0.01332 and 0.777 and 0.01140, respectively. The data based on Cyt *b* gene and D-loop region showed that the genetic diversity of the ST population was lowest among those of the three populations. The genetic distances among the population were 0.00173–0.03615 (Cyt *b*) and 0.00193–0.02639 (D-loop). Obvious genetic differentiation and limited gene exchange were observed between the ST and SL and the ST and LK populations based on fixation indexes ( $F_{st}$ ) and the gene flow value ( $N_m$ ). Analysis of molecular variance (AMOVA) indicated that the source of variation between population (94.60%, 90.69%) was much higher than that within populations (5.40%, 9.31%). The maximum likelihood phylogenetic tree and haplotype network showed that the genetic distance between SL and LK populations was short and formed a branch, while the other branch only consisted of ST haplotypes. A neutral test and mismatch distribution analysis suggested that *P. elongata* has not undergone a recent population expansion. For ST population, neutral test Tajima's  $D$  and Fu's  $F_s$  values were negative ( $P < 0.05$ ), and the mismatch distribution was represented by a single peak curve, which revealed that *P. elongata* populations in the ST area had possibly experienced a population expansion or bottleneck effect during evolution. Overall, *P. elongata* belongs to the category that includes populations with a high haplotype diversity index and low nucleotide diversity index, whereas obvious genetic differentiation among the three populations were observed with the ST population having the lowest genetic diversity. Thus, we recommend that ST populations should be regarded as a management protection unit to prioritize their conservation, while, SL and LK populations should be managed and protected as a whole. Our results could provide scientific evidence for the conservation, sustainable development, and commercial utilization of germplasm resources of *P. elongata*.

**Key words:** *Pseudorasbora elongata*; Cyt *b* gene; D-loop region; genetic diversity; genetic structure

**Corresponding author:** LIAN Yuxi. E-mail: lianyx177@163.com