

基于 SNP 标记的青海湖裸鲤遗传多样性及种群结构研究

罗 慧^{1,2}, 方弟安^{1,2}, 何 苗¹, 毛成诚², 匡 箴², 祁洪芳³, 徐东坡²

1. 上海海洋大学 水产与生命学院, 上海 201306

2. 中国水产科学研究院淡水渔业研究中心/农业农村部淡水渔业和种质资源利用重点实验室, 江苏 无锡 214081

3. 青海湖裸鲤救护中心, 青海 西宁 810000

摘要: 为探究青海湖裸鲤 (*Gymnocypris przewalskii*) 的种质资源现状, 为该物种保护措施的制定提供参考依据。首次利用基因分型技术 (Genotyping-by-sequencing, GBS), 对青海湖水域的6个地理群体共72尾青海湖裸鲤进行单核苷酸多态性 (Single nucleotide polymorphism, SNP) 标记开发和遗传特征分析。共检测出1 600 061个SNP位点, 质控后筛选出45 266个高质量的SNP位点用于遗传分析, 发现核苷酸多样性 (P_i) 为0.317 0~0.327 4, 观测杂合度 (H_o) 和期望杂合度 (H_e) 分别为0.459 4~0.482 3和0.336 7~0.344 4。6个地理群体的遗传距离 (D) 为0.018 4~0.023 3, 两两群体的遗传分化系数 (F_{st}) 均不显著 ($P>0.05$)。分子方差分析 (Analysis of molecular variances, AMOVA) 显示102.37%的遗传变异来自群体内; 群体遗传结构和系统发育进化树分析均显示6个群体属于一个集群, 具有遗传同质性; 而主成分判别分析 (Discriminant analysis of principal components, DAPC) 表明哈尔盖河、黑马河和沙柳河群体相互交叉聚类, 其余3个地理群体分别聚类。综上, 6个青海湖裸鲤群体的 H_o 均大于 H_e , 种群结构单一。

关键词: 青海湖裸鲤; 单核苷酸多态性; 基因分型技术; 遗传分化; 渔业资源

中图分类号: S 932

文献标志码: A

开放科学 (资源服务) 标识码 (OSID):



Genetic diversity and population structure of *Gymnocypris przewalskii* based on SNP markers

LUO Hui^{1,2}, FANG Di'an^{1,2}, HE Miao¹, MAO Chengcheng², KUANG Zhen², QI Hongfang³, XU Dongpo²

1. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China

2. Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences/Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture and Rural Affairs, Wuxi 214081, China

3. Rescues Center of Qinghai Lake Naked Carp, Xining 810000, China

Abstract: To investigate the genetic diversity and genetic differentiation of natural resources of *Gymnocypris przewalskii*, and provide references for its conservation measures, we collected 72 individuals of *G. przewalskii* from six natural populations in Qinghai Lake to analyze the genomic SNP and genetic characteristics by using simplified genome sequencing. Altogether 1 600 061 SNP markers were obtained, and 45 266 high-quality SNP loci were screened for genetic analysis after filtering. The average P_i diversity index was 0.317 0–0.327 4. The average observed heterozygosity (H_o) and expected heterozygosity (H_e) were 0.459 4–0.482 3 and 0.336 7–0.344 4, respectively. The genetic distance (D) was 0.018 4–0.023 3 and the genetic differentiation index (F_{st}) were not significant ($P>0.05$). AMOVA analysis shows that the genetic variation was 102.37% within populations. Further-

收稿日期: 2022-04-02; 修回日期: 2022-05-23

基金项目: 国家重点研发计划项目 (2019YFD0900604); 西北地区重点水域渔业资源与环境调查专项青海湖渔业资源与环境调查 (XB2019-2021); 中国水产科学研究院基本科研业务费专项资金 (2020TD61)

作者简介: 罗 慧 (1998—), 女, 硕士研究生, 研究方向为渔业资源。E-mail: 673131080@qq.com

通信作者: 方弟安 (1978—), 男, 研究员, 博士, 研究方向为渔业资源。E-mail: fangda@ffrc.cn

more, both population structure and phylogenetic tree analysis show that the six *G. przewalskii* populations clustered into one glade with similar genetic structure. In contrast, the genetic structure obtained from DAPC analysis was relatively clearer, showing that the Haergai, Heima and Shaliu River populations cross-clustered with each other, while the remaining three populations clustered into three other branches. In conclusion, the observed heterozygosity of the six geographic populations was higher than the expected value, with a homogeneous population structure.

Keywords: *Gymnocypris przewalskii*; Single nucleotide polymorphism; Genotyping by sequencing; Genetic differentiation; Fishery resources

青海湖裸鲤 (*Gymnocypris przewalskii*) 俗称“湟鱼”, 隶属鲤科、裂腹鱼亚科、裸鲤属, 主要分布于青海湖及周边淡水支流, 为生殖洄游性鱼类, 每年 4—7 月溯河上游至各淡水支流产卵, 生殖行为结束后顺流而下返回青海湖, 是湖中特有的珍稀鱼种, 具有极高的生态价值^[1-2]。由于受到早期的过度捕捞, 附属河道的工程建设, 产卵场、育幼场被破坏, 自身繁殖力低以及气候变化等因素的影响, 青海湖裸鲤的资源量严重衰退^[2]。近年来, 增殖放流及修复洄游通道等工作使其资源量逐渐恢复, 但是对其种质资源的恢复现状未知^[3-4]。目前青海湖裸鲤的研究主要集中在鱼种培育技术^[5-7]、盐度适应性机制^[8-9]、耳石标记^[10-11]、基因的克隆和表达^[12-14]等方面, 关于青海湖裸鲤遗传资源的可用信息非常有限。因此, 为科学保护青海湖裸鲤种质资源, 恢复和提高其资源量, 亟需深入开展种质资源现状评估的相关工作, 尤其是青海湖裸鲤不同地理群体的种质差异和遗传特征。

早期学者分别采用同工酶、限制性片段长度多态性 (RFLP)、扩增片段长度多态性 (AFLP)、线粒体 DNA (mt DNA) 和微卫星 (SSR) 等标记方法研究青海湖裸鲤群体遗传结构和遗传多样性^[15-22]。与上述传统标记相比, 单核苷酸多态性 (Single nucleotide polymorphism, SNP) 作为新一代分子标记, 具有数量多、分布广泛、遗传稳定性高等特点, 且其二等位基因性使其便于自动化检测、适合大样本量分析^[23], 故能在鱼类群体结构及遗传多样性评估中发挥重要作用。随着分子生物学技术的发展, 近年来不断涌现出多种可以获取 SNP 标记的高通量测序方法。其中, 基因分型 (Genotyping by sequencing, GBS) 技术通过酶切加标签的方式, 高通量测序获得可以代表整个基因组信息的部分基因组序列信息, 进而检测大量 SNP 变异位点信息。该方法不受参考基因组限制, 操作过程简单且成本较低, 是进行 SNP 标记开发最为有效、便捷的简化

基因组测序的方法之一^[24]。目前, 已有不少学者利用该技术获取 SNP 标记, 如刘凯等^[25]利用 GBS 技术对三角鲂 (*Megalobrama terminalis*)、翘嘴鲇 (*Culter alburnus*)、蒙古鲇 (*Chanodichthys mongolicus*) 及其杂交子代测序, 获得了 97 911 个 SNP 标记进行遗传结构分析, 为鲂鲇鱼类的杂交选育提供了更多的遗传学数据; Zhang 等^[26]利用 GBS 技术对半滑舌鳎 (*Cynoglossus semilaevis*) 测序获得了 563 109 个 SNP 标记, 并对其遗传多样性水平进行了评估。

到目前为止, 利用 GBS 技术对青海湖裸鲤进行 SNP 标记开发及遗传多样性分析的研究尚未见报道。本研究采用 GBS 测序技术开发青海湖裸鲤的 SNP 位点, 并基于开发出的 SNP 位点对青海湖水域 6 个不同地理群体的青海湖裸鲤进行遗传学分析, 旨在明确青海湖裸鲤的遗传多样性水平、遗传结构及群体间的亲缘关系, 为青海湖裸鲤种质资源的保护和科学管理提供依据。

1 材料与方法

1.1 样品采集

样品采集于青海湖裸鲤繁殖洄游期 (2019 年和 2020 年的 5—9 月), 采用渔业资源底拖网开展样品采集, 分别采集于青海湖附近 5 条主要产卵河道 [布哈河 (BHH)、哈尔盖河 (HEG)、黑马河 (HMH)、泉吉河 (QJH)、沙柳河 (SLH)] 和青海湖湖区一郎剑 (YLJ) 共 6 个不同地点 (图 1, 表 1)。无侵入性伤害剪去部分鳍条, 并保存于无水乙醇中。采用 ArcGIS 10.2 软件绘制地图。

1.2 实验方法

1.2.1 DNA 提取与 GBS 文库构建

DNA 提取: 使用 CTAB (十六烷基三甲基溴化铵) 法提取样品基因组 DNA, 送广州基迪奥生物科技有限公司进行 DNA 质检、建库和测序。利用质量分数为 1.0% 的琼脂糖凝胶电泳、Qubit 荧光光

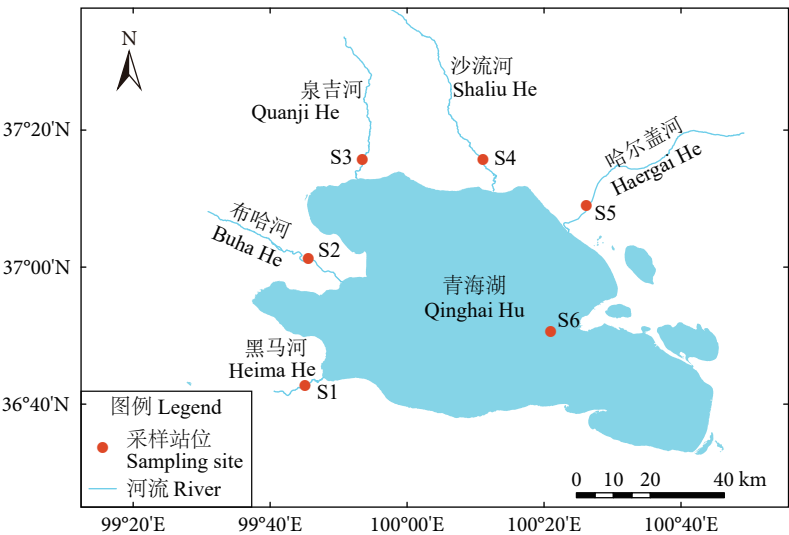


图1 青海湖裸鲤采样地点图

Fig. 1 Sampling site of *G. przewalskii*

表1 青海湖裸鲤的样品采集信息

Table 1 Sample information of *G. przewalskii*

采集地点 Sampling site	经纬度 Latitude and longitude	数量 Sample size/尾	取样部位 Sampling part
布哈河 BHH	99°44'24"E, 37°02'24"N	12	尾鳍
哈尔盖河 HEG	100°28'48"E, 37°13'12"N	12	尾鳍
黑马河 HMH	99°46'12"E, 36°43'48"N	12	尾鳍
泉吉河 QJH	99°53'24"E, 37°14'24"N	12	尾鳍
沙柳河 SLH	100°10'48"E, 37°14'24"N	12	尾鳍
一郎剑 YLJ	100°24'00"E, 36°42'00"N	12	尾鳍

度计 (Thermo Fisher Scientific, Waltham, MA) 和 NanoDrop 2000 分光光度计 (Thermo Fisher Scientific, Waltham, MA) 检测 DNA 的完整性、总量和纯度。选择合格的 DNA 用于 GBS 文库构建和测序。

GBS 文库构建：首先应用限制性内切酶 *Eco*RI 和 *Nla*III 对基因组进行酶切，酶切后的片段加上带有 barcode 的接头后，经过 AMPure XP 磁珠进行纯化，并选择 300~400 bp 范围的片段进行 PCR 扩增，建好的文库混合后经过纯化、库检。

1.2.2 测序和数据质控

利用 Illumina Hiseq 测序平台，进行双末端 (Paired-End) 150 测序。对原始下机数据 (Raw data) 进行质控过滤，质控标准为：1) 过滤掉含有接头序列的短读序列；2) 去除含有未知核苷酸 $\geq 10\%$ 的序列；3) 单端测序中低质量 (质量值 $Q \leq$

10) 碱基数超过该条序列长度 50% 的序列。经过对测序数据的严格质控，得到高质量的测序序列。

1.3 数据分析

1.3.1 SNP 标记开发

对于数据质控获得的高质量序列，利用 Stacks 软件^[27] 进行聚类。聚类后的标签进行去重处理，没有重复的标签用 N 进行链接，然后用脚本制作成大小一致的参考基因组进行后续分析。利用 BWA^[28] 将测序序列比对到参考基因组上，使用软件 GATK (3.4-46)^[29] 进行多个样本的变异检测，检测到的 SNP 位点使用 Variant Filtration 程序进行过滤。

1.3.2 群体遗传多样性分析

对获取的 SNP 位点进行筛选，筛选标准为：1) 测序深度 (Depth) ≥ 5 ，即当测序深度 < 4 时，对应位点基因型矫正为缺失；2) 数据完整度 $\geq 80\%$ ；3) 次要基因型频率 (MAF) ≥ 0.05 ；4) 保留双等位基因。筛选后，使用符合条件的 SNP 用于青海湖裸鲤群体遗传多样性和遗传结构分析。

为评价青海湖裸鲤 6 个地理群体的遗传多样性水平，用 Genepop 软件^[30] 计算每个 SNP 位点的观测杂合度 (Observed heterozygosity, H_o)、观测纯合度 (Observed homozygosity, OH)、期望杂合度 (Expected heterozygosity, H_e)、期望纯合度 (Expected homozygosity, EH)；用 Popgenome 软件包^[31] 计算位点的核苷酸多样性指数 (Nucleotide diversity degree, P_i)；再对各群体中每个位点的指数求和，然后求均值。

为了解遗传变异来源和种群分化情况, 使用 Arlequin 3.5.2.2 软件^[32] 基于 Pairwise differences 方法进行分子方差分析 (Analysis of molecular variance, AMOVA), 并计算两两群体间的遗传分化指数 (F_{st}), 显著性水平通过 10 000 次重抽样模拟 (Bootstrapping permutations) 进行计算; 用 MEGA 11.0.10 软件^[33] 中的 Kimura 2-parameter model 模型计算群体间的遗传距离 (D)。

利用基于最大似然法 (Maximum likelihood estimation) 的聚类软件 Admixture^[34] 解析群体遗传结构, 确定最佳分群情况, 选取 $K=1\sim7$ 进行计算, 并进行交叉验证, 选取交叉验证错误率最小值时的 K 值作为最佳 K 值; 为分析个体间的进化关系, 用软件 Treebest^[35] 邻接法 (Neighbor-joining methods) 基于 P-distance 构建进化树, 随后用 iTOL 软件^[36] 对进化树进行美化; 使用 R 语言的 Adegenet 软件包^[37] (参数: $n.da=5$, $npca=33$) 做主成分判别分析 (Discriminant analysis of principal components, DAPC)。

2 结果

2.1 测序数据及 SNP 检测统计

GBS 测序共得到测序数据 194.94 Gb, 经数据质控过滤后, 得到高质量 reads 1 369 672 960 条, 高质量数据 192.29 Gb, 平均每个样品数据量 2.67 Gb。质控后的测序结果具有较高的质量 ($Q_{20}\geq 96.75\%$, $Q_{30}\geq 91.16\%$) 和稳定的 GC 含量 ($38.17\%\sim 39.92\%$), 数据质量满足分析需要 (表 2)。Reads 聚类后得到 SNP 位点 1 600 061 个, 得到标签 531 183 个。根据选择标准, 进一步筛选、过滤后获得 45 266 个高质量的 SNP 位点用于群体遗传分析。

2.2 遗传多样性分析

遗传多样性参数的统计结果发现, 6 个青海湖裸鲤群体的 H_o 为 0.459 4~0.482 3, H_e 为 0.336 7~0.344 4, P_i 为 0.317 0~0.327 4 (表 3)。6 个青海湖裸鲤群体的 H_o 均高于 H_e , 表明实际的等位基因多态性高于预测。在 6 个群体中, 泉吉河群体的 H_o 、 H_e 、 P_i 均最大, 一郎剑群体的 H_o 、 H_e 、 P_i 均最小。

2.3 遗传分化分析

2.3.1 遗传距离与遗传分化系数

计算得出所有群体成对的 D 和 F_{st} 矩阵 (表 4), 6 个地理群体间的 D 介于 0.018 4~0.023 3, F_{st} 介于 $-0.029\ 8\sim -0.021\ 2$, 且均不显著 ($P>0.05$)。表明

6 个群体之间遗传距离近, 在遗传上不存在显著性差异。

2.3.2 群体间分子变异分析

AMOVA 分析结果显示 6 个青海湖裸鲤地理群体间未形成显著的遗传分化 ($P>0.05$), 且遗传变异主要来源于群体内 (表 5)。

2.4 遗传结构分析

2.4.1 群体结构分析

群体遗传结构分析见图 2, 交叉验证错误率的最小值是 $K=1$ 时, 表明 6 个地理群体的青海湖裸鲤未形成明显的谱系结构, 为一个种群 (图 3)。

2.4.2 系统发育进化树分析

从进化树上可以得出来自 6 个不同地理群体的青海湖裸鲤样本存在明显的混杂情况, 即同一地理群体样本并未完全聚在一起, 甚至被分到不同的分支, 说明青海湖裸鲤样本间并不完全按照地理分布聚类, 存在多歧分支 (图 4)。

2.4.3 DAPC 分析

DAPC 分析结果揭示了青海湖裸鲤更为清晰的群体间遗传结构。DAPC 分析基于最优的主成分数 (最优主成分数 $PC=33$) 对不同青海湖裸鲤群体进行聚类分析, 分析结果显示不同采样位点的青海湖裸鲤群体划分为 4 组 (图 5)。虽然不同组之间存在一定程度的重叠, 但基于分析结果可以看出哈尔盖河、黑马河和沙柳河的青海湖裸鲤群体主要聚为 1 组, 布哈河、泉吉河、一郎剑分别各自聚为另外 3 组 (图 6)。

3 讨论

3.1 青海湖裸鲤的遗传多样性

遗传多样性是评估种群资源状况的重要依据, 与物种的生存、适应能力及进化潜能密切相关。杂合度是度量群体遗传多样性的关键指标之一, 其数值越大, 群体变异程度越高, 遗传多样性越丰富^[17]。Jiang 等^[38] 利用飞行质谱阵列 SNP 基因分型技术 (Sequenom MassARRAY) 分析了 3 个大黄鱼 (*Larimichthys crocea*) 群体的遗传多样性, 发现 H_o 为 0.190~0.253, H_e 为 0.273~0.320; 李光华等^[39] 利用开发的 598 354 个 SNP 标记对 20 尾短须裂腹鱼 (*Schizothorax wangchiachii*) 进行群体遗传多样性分析, 发现为 H_o 为 0.200 7, H_e 为 0.316 0; 金方彭等^[40] 对同为裂腹鱼亚科的澜沧江 4 个光唇裂腹鱼 (*S. lisolabiatus* Tsao) 地理群体进行了简化基因组测序,

表2 GBS测序数据统计
Table 2 Data statistics after genotyping by sequencing

样品 Sample	原始数据量 Raw base/bp	有效测序数量 Clean base/bp	Q20/%	Q30/%	比对率 Ratio/%	覆盖度 Cover/%	GC/%
BHH01	1 389 577 920	1 360 512 534	97.20	92.24	72.38	41.86	39.43
BHH02	1 927 663 064	1 910 163 826	97.00	91.73	71.71	44.45	39.28
BHH03	1 355 388 864	1 330 014 564	97.23	92.29	72.40	41.63	39.38
BHH04	1 854 567 996	1 840 004 206	97.14	92.08	68.34	44.96	39.35
BHH05	2 265 476 520	2 228 411 786	97.24	92.33	74.30	44.36	39.34
BHH06	1 602 581 328	1 578 477 056	97.16	92.14	72.72	43.72	39.33
BHH07	1 493 156 128	1 486 185 014	96.56	90.65	68.29	41.88	39.62
BHH08	3 625 787 104	3 582 205 358	96.78	91.19	69.32	47.85	39.82
BHH09	4 188 671 696	4 152 405 788	96.70	91.03	66.88	49.40	39.92
BHH10	3 783 676 336	3 743 104 234	96.83	91.29	69.96	48.18	39.83
BHH11	2 886 390 548	2 857 277 330	96.82	91.26	69.67	48.13	39.83
BHH12	2 436 548 180	2 404 222 590	96.86	91.36	71.23	45.73	39.80
HEG01	4 452 147 016	4 387 873 466	96.49	90.46	71.07	50.54	39.53
HEG02	3 773 555 144	3 725 723 130	96.44	90.34	70.17	48.07	39.55
HEG03	4 206 118 100	4 143 796 870	96.45	90.39	71.59	48.20	39.51
HEG04	5 078 768 024	5 001 034 014	96.47	90.42	71.66	49.81	39.56
HEG05	1 153 394 716	1 139 883 590	96.54	90.61	73.61	39.84	39.45
HEG06	4 843 459 540	4 792 194 098	96.43	90.37	73.01	48.96	39.54
HEG07	1 501 577 864	1 481 599 170	97.26	92.36	74.21	40.65	39.47
HEG08	2 693 727 504	2 657 441 746	97.31	92.46	76.19	44.39	39.39
HEG09	1 194 213 752	1 168 489 596	97.31	92.39	76.99	38.43	39.40
HEG10	1 372 619 144	1 339 243 546	96.89	91.50	72.82	40.60	39.33
HEG11	3 088 083 656	3 050 778 984	97.11	92.02	74.10	45.97	39.42
HEG12	1 619 334 204	1 585 556 642	97.16	92.14	75.77	41.52	39.28
HMH01	1 232 127 752	1 200 318 504	97.02	91.83	71.56	40.28	39.22
HMH02	3 026 290 652	2 985 218 180	97.23	92.31	73.98	44.96	39.25
HMH03	1 399 119 468	1 372 569 888	97.27	92.41	73.46	41.05	39.17
HMH04	1 970 927 056	1 921 307 060	97.28	92.39	73.24	43.12	39.30
HMH05	2 647 884 508	2 596 362 030	97.23	92.30	74.44	44.43	39.30
HMH06	1 708 154 352	1 660 462 320	97.29	92.44	73.06	41.96	39.26
HMH07	4 025 817 860	3 985 350 084	97.69	93.28	70.09	48.71	38.17
HMH08	5 205 970 488	5 117 137 470	96.65	90.89	73.81	48.98	39.20
HMH09	4 940 601 456	4 863 779 832	96.65	90.91	75.28	48.08	39.24
HMH10	1 963 437 692	1 945 709 624	96.60	90.76	73.03	43.03	39.22
HMH11	1 666 515 692	1 642 075 532	96.55	90.67	72.54	41.39	39.09
HMH12	4 347 827 000	4 286 730 938	96.56	90.69	71.40	48.89	39.35

续表 2

to be continued

样品 Sample	原始数据量 Raw base/bp	有效测序数量 Clean base/bp	Q20/%	Q30/%	比对率 Ratio/%	覆盖度 Cover/%	GC/%
QJH01	5 147 123 416	5 090 638 928	96.32	90.09	70.17	55.23	39.50
QJH02	3 584 795 680	3 565 628 068	95.95	89.37	66.88	49.55	39.44
QJH03	6 454 315 492	6 410 786 704	95.95	89.35	71.12	51.56	39.43
QJH04	3 914 473 956	3 894 271 390	95.87	89.22	67.09	48.68	39.41
QJH05	5 765 227 264	5 711 763 208	96.41	90.36	73.91	60.43	39.40
QJH06	1 578 283 708	1 552 126 286	96.41	90.24	74.89	50.12	39.39
QJH07	1 396 840 084	1 358 462 200	96.85	91.53	72.06	47.11	39.26
QJH08	2 275 133 940	2 251 308 686	97.01	91.90	70.44	52.56	39.29
QJH09	1 340 744 972	1 315 726 238	97.13	92.17	74.67	41.01	39.25
QJH10	2 048 956 056	2 009 990 568	97.09	92.05	71.84	44.26	39.38
QJH11	2 092 240 212	2 061 253 446	97.04	91.95	71.65	45.25	39.40
QJH12	2 039 776 892	1 992 244 374	97.13	92.16	72.73	44.70	39.35
SLH01	7 040 469 056	6 945 862 130	96.65	90.81	70.87	52.34	39.45
SLH02	4 498 263 220	4 427 530 468	96.69	90.89	71.71	49.57	39.50
SLH03	5 164 724 600	5 075 426 948	96.66	90.86	71.14	50.79	39.57
SLH04	1 272 379 640	1 259 999 052	96.63	90.74	73.42	39.66	39.41
SLH05	1 050 489 020	1 032 232 476	96.59	90.67	73.08	37.98	39.22
SLH06	2 360 256 124	2 322 023 500	96.63	90.75	72.68	44.17	39.56
SLH07	2 787 269 152	2 752 924 192	97.26	92.34	71.90	45.18	39.52
SLH08	1 659 152 708	1 635 202 024	97.22	92.24	70.67	42.13	39.55
SLH09	1 974 618 488	1 936 922 806	97.16	92.16	68.63	43.39	39.57
SLH10	1 434 149 732	1 414 000 878	97.20	92.20	67.84	42.37	39.50
SLH11	1 465 104 312	1 452 041 076	97.13	92.12	67.80	42.89	39.64
SLH12	1 593 403 584	1 565 068 850	96.44	90.59	73.79	41.90	39.02
YLJ01	1 458 951 168	1 441 355 132	97.14	92.12	71.91	41.68	39.51
YLJ02	1 933 206 744	1 917 865 930	97.14	92.17	67.47	44.98	39.14
YLJ03	3 128 604 208	3 093 206 580	97.14	92.21	72.61	47.09	39.06
YLJ04	1 343 484 720	1 316 565 174	97.24	92.32	76.23	40.09	38.99
YLJ05	1 588 695 432	1 551 573 180	97.07	91.93	72.37	42.47	38.97
YLJ06	3 330 149 352	3 289 825 750	97.25	92.34	70.46	47.58	39.00
YLJ07	1 873 001 868	1 844 675 484	97.30	92.47	72.27	43.94	39.03
YLJ08	2 175 175 880	2 142 504 822	96.39	90.46	71.81	44.84	39.17
YLJ09	3 172 296 188	3 134 366 418	96.33	90.34	72.60	47.36	39.14
YLJ10	2 801 743 496	2 774 569 032	96.34	90.37	69.66	47.28	39.06
YLJ11	2 270 060 848	2 236 343 778	96.47	90.63	72.99	44.74	39.09
YLJ12	2 007 127 400	1 987 101 014	96.34	90.38	71.46	44.26	39.08

表3 青海湖裸鲤遗传多样性分析
Table 3 Genetic diversity of *G. przewalskii*

群体 Population	观测杂合度 H_o	观测纯合度 OH	期望杂合度 H_e	期望纯合度 EH	核苷酸多样性 P_i
布哈河 BHH	0.459 8	0.540 2	0.337 3	0.662 7	0.317 2
哈尔盖河 HEG	0.468 6	0.531 4	0.342 1	0.657 9	0.321 7
黑马河 HMH	0.466 1	0.533 9	0.340 0	0.660 0	0.320 5
泉吉河 QJH	0.482 3	0.517 7	0.344 4	0.655 6	0.327 4
沙柳河 SLH	0.465 8	0.534 2	0.343 3	0.656 8	0.318 1
一郎剑 YLJ	0.459 4	0.540 6	0.336 7	0.663 3	0.317 0

表4 青海湖裸鲤的遗传距离和遗传分化系数表
Table 4 Genetic differentiation and distance of *G. przewalskii*

群体 Population	BHH	HEG	HMH	QJH	SLH	YLJ
布哈河 BHH		0.022 8	0.022 7	0.019 1	0.023 2	0.023 3
哈尔盖河 HEG	-0.021 7		0.021 9	0.018 6	0.022 6	0.022 7
黑马河 HMH	-0.022 3	-0.022 4		0.018 4	0.022 5	0.022 4
泉吉河 QJH	-0.021 5	-0.023 0	-0.023 0		0.019 2	0.019 3
沙柳河 SLH	-0.027 2	-0.025 7	-0.026 6	-0.029 8		0.023 2
一郎剑 YLJ	-0.021 4	-0.021 7	-0.022 4	-0.021 2	-0.027 0	

注：下三角为群体间遗传分化系数，上三角为群体间遗传距离。

Note: The lower triangle is the genetic differentiation coefficient (F_{st}) among populations, and the upper triangle is the genetic distance (D) among populations.

表5 6个青海湖裸鲤群体的分子变异分析 (AMOVA)
Table 5 Analysis of molecular variation (AMOVA) of six populations of *G. przewalskii*

变异来源 Source of variation	自由度 Degree of freedom	平方和 Sum of squares	变异组分 Variance components	变异百分率 Percentage of variation/%
群体间 Among population	5	13 287.29	-138.528 8Va	-2.37
群体内 Within population	138	825 536.29	5 982.147 0Vb	102.37
总计 Total	143	838 823.58	5 843.618 3	

发现 H_o 为 0.238 1~0.313 1, H_e 为 0.269 0~0.311 5。从上述结果来看, 与同样以 SNP 为分子标记的其他鱼类遗传多样性相比较, 本研究的杂合度水平 (H_o : 0.459 4~0.482 3, H_e : 0.336 7~0.344 4) 相对较高, 且 6 个群体的 H_o 均大于 H_e , 呈现杂合子过剩现象。此外, 本研究结果还与之前基于微卫星标记的青海湖裸鲤遗传多样性研究结果相似, 蒋鹏等^[17] 用 10 个多态的微卫星分子标记对布哈河、沙柳河、哈尔盖河、泉吉河、黑马河和淡水系的 6 个青海湖裸鲤群体遗传多样性进行研究, 6 个群体杂合度水平 (H_o : 0.414 5~0.510 5) 较高, 具有一定的遗传变异潜力。

P_i 是衡量群体遗传多样性的另一重要指标,

P_i 值越高表明遗传多样性越丰富。由于分子标记方法的不同, 本研究获得的 P_i 值与之前基于线粒体控制区序列的青海湖裸鲤的 P_i 值相差两个数量级, 故无法进行直接比较^[16,19,22]。

3.2 青海湖裸鲤无遗传分化

鱼类群体之间的 D 和 F_{st} 是反映群体分化程度的主要指标。Shaklee 等^[41] 提出鱼类在属、种和种群三级水平上的 D 值分别为 0.9、0.30 和 0.05, 本研究中 D 值为 0.018 4~0.023 3, 表明 6 个青海湖裸鲤群体属同一种群, 亲缘关系近, 这与蒋鹏等^[17] 的研究结果相近。Wright^[42] 提出: $F_{st}<0.05$ 群体间呈低度分化; $0.05<F_{st}<0.15$ 呈中度分化; $0.15<F_{st}<0.25$ 分化程度较大; $F_{st}>0.25$ 分化程度极大。

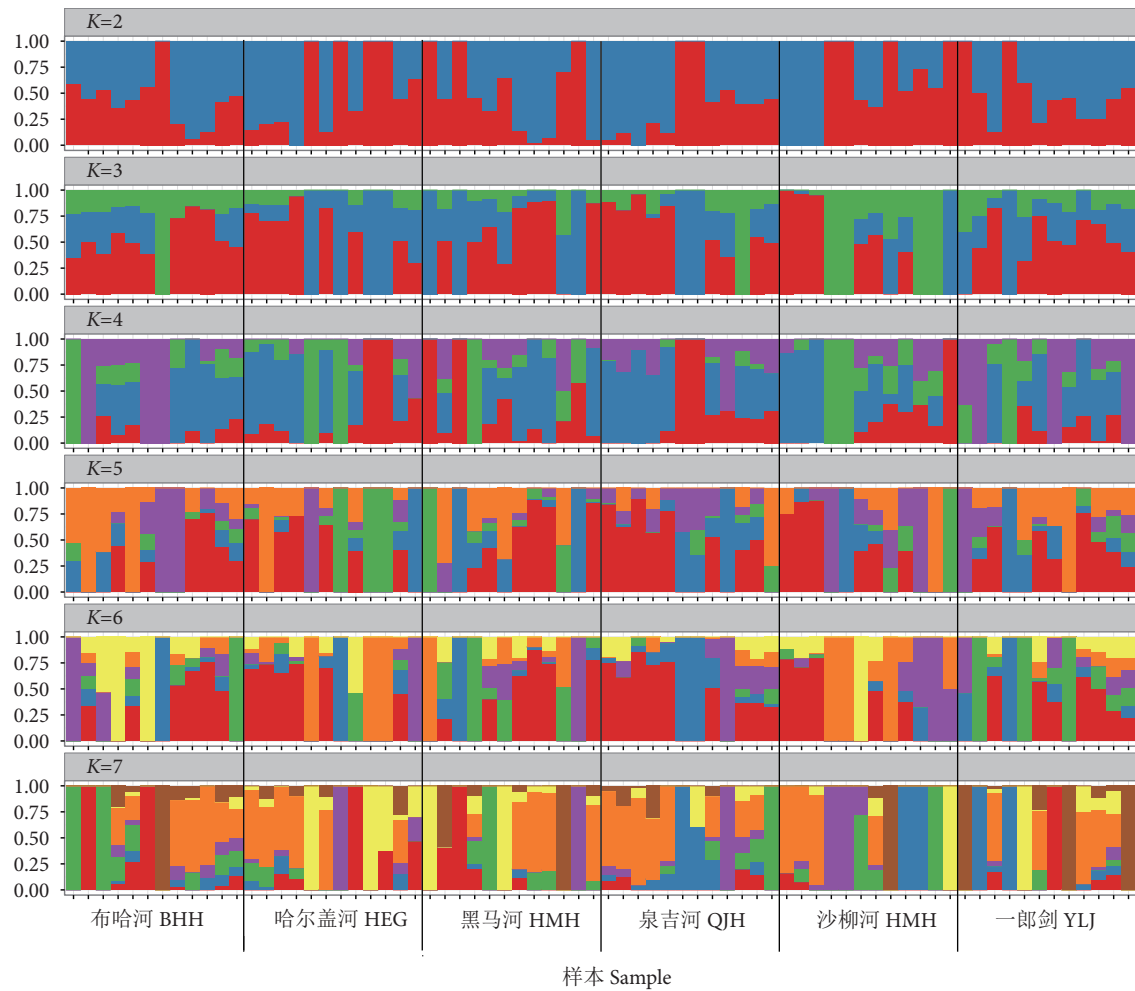


图2 青海湖裸鲤样品聚类结果

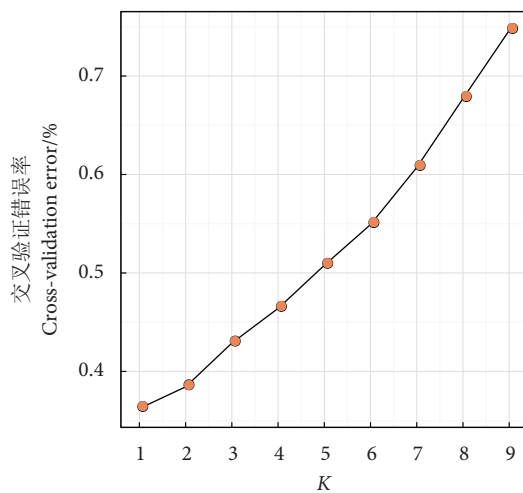
Fig. 2 Clustering results of *G. przewalskii* samples

图3 ADMIXTURE 交叉验证K值曲线图

Fig. 3 ADMIXTURE cross-validation K value plot

本研究中 6 个不同地理群体间 F_{st} 值最大为 -0.0212 ($P>0.05$), 表明不同地理群体间并不存在显著的遗传分化。AMOVA 的结果显示 102.37% 以上的遗传变异来自群体内, 也说明了群体间无遗传分化, 这

与赵凯等^[43]的研究结果一致。陈大庆等^[19]和张春霖等^[44]分别利用 D-loop 和 RAPD 的方法对青海湖附属的 3 条河流 (黑马河、布哈河和沙柳河) 的裸鲤遗传多样性进行研究, 结果表明遗传变异主要存在于繁殖群体内, 繁殖群体间的基因变异较弱, 洄游到同一河流里进行交配繁殖的群体内基因交流作用比较大, 而洄游到不同河流进行繁殖的群体间的基因交流相对较小。张仁意等^[16]对青海湖裸鲤 4 个地理群体 (青海湖、可鲁克湖、甘子河、草搭连) 的遗传多样性进行研究发现, 种群间出现了一定程度的遗传分化, 特别是可鲁克湖种群已经高度分化。推测上述研究结果存在较大差异的主要原因是检测手段和样本来源不同, 而本研究中 GBS 技术所选用的遗传标记数量多, 能够覆盖全基因组, 因此具有更高的可靠性。

3.3 青海湖裸鲤的遗传结构特征

群体结构和系统发育进化树分析的结果进一步验证了上述遗传分化分析的结果, 均把 6 个青海湖

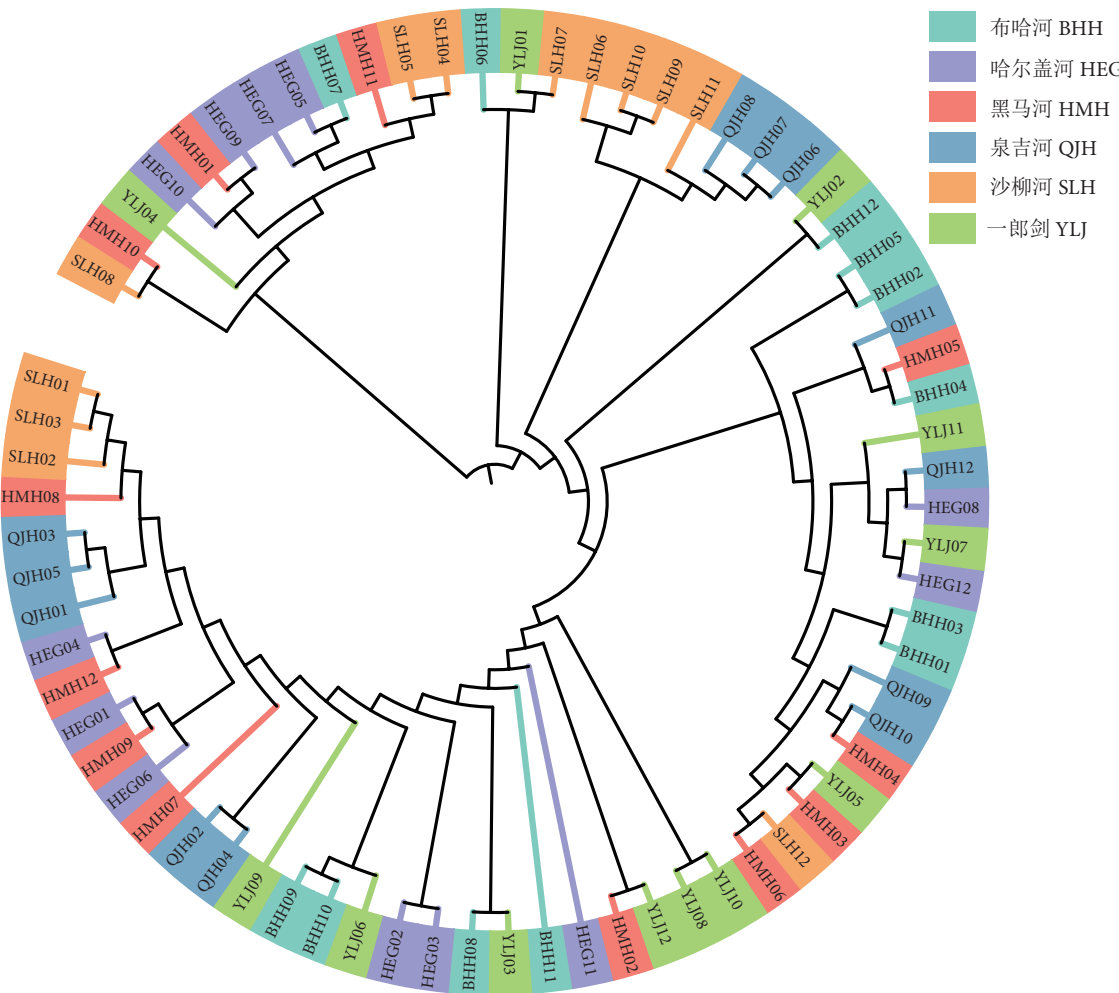


图4 6个青海湖裸鲤群体的系统进化树
Fig. 4 Phylogenetic tree of six *G. przewalskii* populations

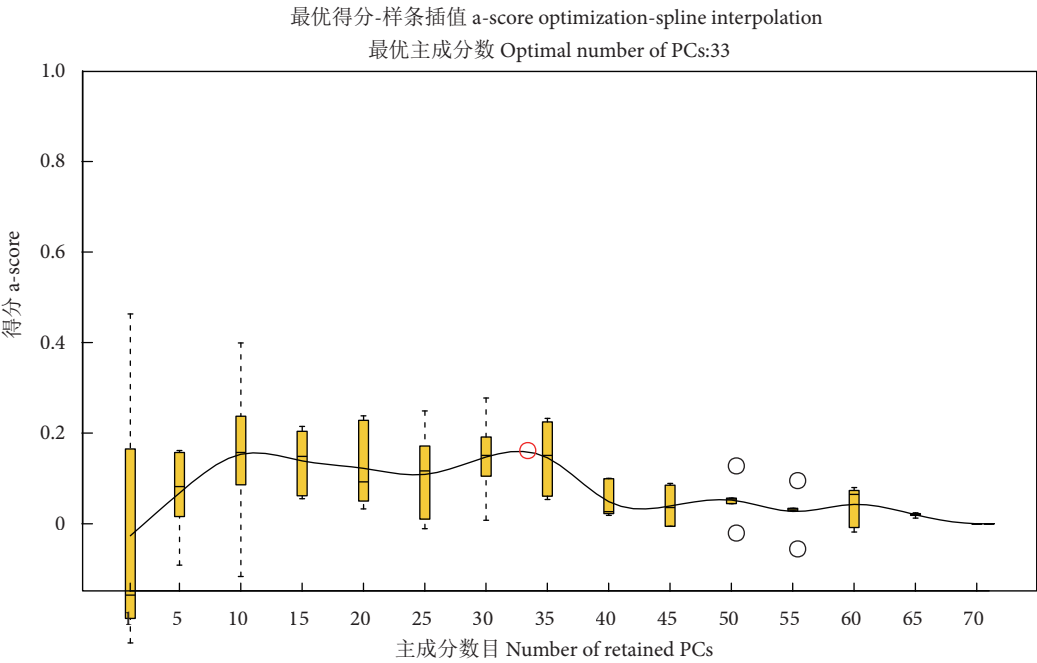
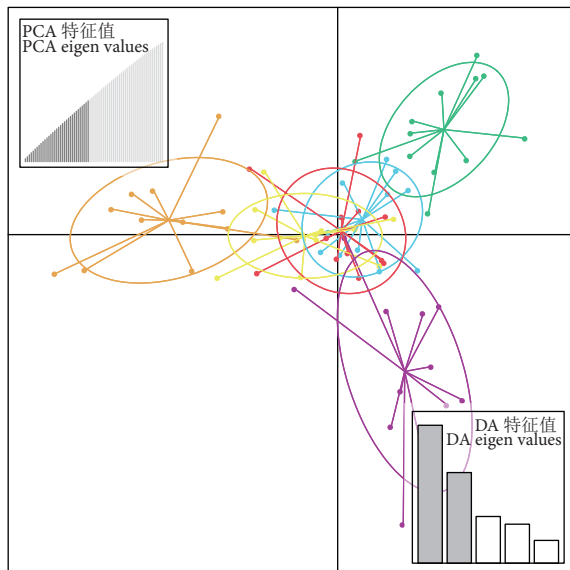


图5 DAPC最优主成分数目的a-score优化
Fig. 5 DAPC a-score optimization of number of retained PCs



每个点代表 1 个青海湖裸鲤个体, 每个椭圆代表 1 个群体。
Each point represents a *G. przewalskii* individual and each ellipse represents a group.

图6 DAPC聚类图

Fig. 6 Plots of DAPC clusters

裸鲤群体划分为 1 个聚类群体, 表明 6 个群体间不存在显著的遗传差异。然而, DAPC 分析方法获得了相对更加清晰的群体遗传结构。DAPC 的结果显示, 6 个青海湖裸鲤群体主要聚为 4 组, 其中, 哈尔盖河、黑马河和沙柳河群体聚在一起, 另外 3 个群体各自聚为 3 组。推测哈尔盖河和沙柳河群体聚在一起可能是因为地理距离较近, 相互间基因交流频繁, 而黑马河可能是因为生境变化、种群瓶颈效应和人工增殖放流等影响了其群体遗传结构^[17]。此外, 泉吉河和布哈河分别聚为两组, 可能是因为泉吉河和布哈河是两条较大的入湖河流, 种群数量大而稳定, 遗传结构还保存着各自的特点; 也可能与青海湖裸鲤的生殖洄游有关, 早期的一些研究表明青海湖裸鲤已经建立了一种相对稳定的生殖洄游模式^[17,45-46]。一郎剑群体为湖中群体, 独自聚为一组, 可能与地理位置相关, 其采样点位于青海湖东南方且附近无入湖河流, 该群体可能未参加生殖洄游。

3.4 青海湖裸鲤种质资源的保护建议

近年来, 为恢复青海湖裸鲤的种质资源量, 每年都开展青海湖裸鲤增殖放流活动, 但其对野生种群结构的影响不容忽视。在实际调查中发现, 青海湖裸鲤救护中心进行增殖放流工作时所选用的野生亲本基本都来源于泉吉河和沙柳河, 长期以往必将影响青海湖裸鲤的遗传结构^[15]。因此, 建立科学

合理的人工增殖放流体系势在必行, 在增殖放流过程中应多河道引进亲本并定期扩增或更换亲本, 持续监测青海湖裸鲤的种质资源状况, 以维持青海湖裸鲤丰富的遗传多样性。有研究表明, 迁地保护也是保护物种多样性的一种有效措施^[47], 因此, 可以尝试在地理隔离形成淡水陆封群体中 (如与青海湖毗邻的淡水湖尕斯库勒湖), 开展青海湖裸鲤纯种引入培育, 从而改善其资源状况, 提高青海湖裸鲤的种质资源多样性水平。

综上, 本研究首次利用 GBS 技术成功开发了 6 个青海湖裸鲤地理群体的 SNP 位点, 并对其遗传多样性水平和遗传结构特征进行了分析, 发现 6 个地理群体的观测杂合度均高于期望杂合度, 未形成显著的种群遗传分化, 为单一种群结构。后续可扩大样本量和研究范围, 进一步丰富和完善青海湖裸鲤的遗传标记资源。本研究阐明了青海湖裸鲤的种质资源现状, 为科学保护措施制定提供了基础资料。

参考文献:

- [1] 史建全. 青海湖裸鲤研究现状与资源保护对策 [J]. 青海科技, 2008, 15(5): 13-16.
- [2] 汪松, 解焱. 中国物种红色名录 [M]. 北京: 高等教育出版社, 2004: 468.
- [3] HITOSHI A, CORINNE S. Is hatchery stocking a help or harm? [J]. Aquaculture, 2010, 308(S1): S2-S11.
- [4] 史建全, 祁洪芳, 杨建新, 等. 青海湖裸鲤增殖放流效果评估 [J]. 农技服务, 2016, 33(12): 128-129.
- [5] 丰朝晖, 俞录贤, 付生云, 等. 青海湖裸鲤池塘“春片”鱼种的培育措施 [J]. 科学养鱼, 2021(3): 11-12.
- [6] 吴艳红, 江华敏, 付生云. 工厂化循环水车间青海湖裸鲤商品鱼养殖技术 [J]. 科学养鱼, 2020(4): 82-83.
- [7] 丰朝晖, 俞录贤, 罗颖, 等. 青海湖裸鲤鱼种培育水质调节技术 [J]. 科学养鱼, 2019(5): 7-9.
- [8] 黄岫, 李长忠, 李梓瑄, 等. 盐度对青海湖裸鲤生长及渗透调节基因的影响 [J]. 水产科学, 2022, 41(4): 527-536.
- [9] 王萍, 来琦芳, 么宗利, 等. 盐碱环境下青海湖裸鲤肠道 HCO_3^- 分泌相关基因表达差异 [J]. 海洋渔业, 2015, 37(4): 341-348.
- [10] 江华敏, 吴金平, 吴艳红, 等. 茜素络合物对青海湖裸鲤幼鱼的耳石标记研究 [J]. 科学养鱼, 2021(3): 72-74.
- [11] 周玲, 金章东, WILLIAMS I S, 等. 青海湖裸鲤耳石轮纹 O-Sr 同位素组成及其对洄游行为的指示 [J]. 科学通报, 2016, 61(6): 668-675.
- [12] 周其椿, 张显波, 李建光, 等. 青海湖裸鲤 *eif5b* 基因克隆与初步功能分析 [J]. 贵州农业科学, 2021, 49(2): 92-96.
- [13] 马清花, 陈雪妍, 卫唯, 等. 青海湖裸鲤 *AP-1* 基因的克隆与表达分析 [J]. 基因组学与应用生物学, 2020, 39(7): 2964-2971.

- [14] 梁健, 陈雪妍, 卫唯, 等. 青海湖裸鲤 *TOBI* 和 *TOB2* 基因的克隆与表达分析 [J]. 西北农林科技大学学报 (自然科学版), 2020, 48(5): 31-41.
- [15] 王朝溪, 史建全, 卫福磊, 等. 青海湖不同支流中青海湖裸鲤的 AFLP 遗传多样性分析 [J]. 基因组学与应用生物学, 2015, 34(1): 82-89.
- [16] 张仁意, 李国刚, 汤永涛, 等. 青海湖裸鲤线粒体 DNA D-loop 区的遗传多样性及其遗传分化研究 [J]. 中国农学通报, 2013, 29(32): 71-76.
- [17] 蒋鹏, 史建全, 张研, 等. 应用微卫星多态分析青海湖裸鲤 (*Gymnocypris przewalski* (Kessler)) 六个野生群体的遗传多样性 [J]. 生态学报, 2009, 29(2): 939-945.
- [18] 孟鹏, 王伟继, 孔杰, 等. 五条河流青海湖裸鲤的同工酶变异 [J]. 动物学报, 2007(5): 892-898.
- [19] 陈大庆, 张春霖, 鲁成, 等. 青海湖裸鲤繁殖群体线粒体基因组 D-loop 区序列多态性 [J]. 中国水产科学, 2006, 13(5): 800-806.
- [20] 许生成, 李太平, 李均祥, 等. 青海湖裸鲤线粒体 DNA 多态性研究 [J]. 黑龙江畜牧兽医, 2003(1): 11-12.
- [21] 祁得林. 青海湖裸鲤遗传多样性研究 [D]. 杭州: 浙江大学, 2002: 51.
- [22] 赵凯, 李军祥, 张亚平, 等. 青海湖裸鲤 mtDNA 遗传多样性的初步研究 [J]. 遗传, 2001, 23(5): 445-448.
- [23] VIGNAL A, MILAN D, SANCRISTOBAL M, et al. A review on SNP and other types of molecular markers and their use in animal genetics[J]. Genet Sel Evol, 2002, 34(3): 275-305.
- [24] ELSHIRE R J, GLAUBITZ J C, SUN Q, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species[J]. PLoS One, 2017, 6(5): e19379.
- [25] 刘凯, 冯晓宇, 沈玉帮, 等. 基于 GBS 技术分析鲢鳙鱼类及其杂交子代的遗传结构 [J]. 水产学报, 2021, 45(8): 1307-1316.
- [26] ZHANG B, JIA L, HE X, et al. Large scale SNP unearthing and genetic architecture analysis in sea-captured and cultured populations of *Cynoglossus semilaevis*[J]. Genomics, 2020, 112(5): 3238-3246.
- [27] ROCHETTE N C, CATCHEN J M. Deriving genotypes from RAD-seq short-read data using Stacks[J]. Nat Protoc, 2017, 12(12): 2640-2659.
- [28] DURBIN L R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. Bioinformatics, 2010, 25(14): 1754-1760.
- [29] van der AUWERA G A, CARNEIRO M O, HARTL C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline[J]. Curr Protoc Bioinformatics, 2013, 43(1110): 10-11.
- [30] ROUSSET F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux[J]. Mol Ecol Resour, 2008, 8(1): 103-106.
- [31] PFEIFER B, WITTELSBÜRGER U, RAMOS-ONSINS S E, et al. PopGenome: an efficient Swiss army knife for population genomic analyses in R[J]. Mol Biol Evol, 2014, 31(7): 1929-36.
- [32] LAURENT E, GUILLAUME L, STEFAN S. Arlequin (version 3.0): an integrated software package for population genetics data analysis[J]. Evol Bioinform, 2007, 1: 47-50.
- [33] SUDHIR K, GLEN S, MICHAEL L, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms[J]. Mol Biol Evol, 2018, 35(6): 1547-1549.
- [34] ALEXANDER D H, NOVEMBRE J, LANGE K. Fast model-based estimation of ancestry in unrelated individuals[J]. Genome Res, 2009, 19(9): 1655-1664.
- [35] VILELLA A J, SEVERIN J, URETA-VIDAL A, et al. EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates[J]. Genome Res, 2009, 19(2): 327-335.
- [36] IVICA L, PEER B. Interactive Tree Of Life (iTOL) v4: recent updates and new developments[J]. Nucleic Acids Res, 2019, 47(W1): W256-W259.
- [37] THIBAUT J, SÉBASTIEN D, FRANÇOIS B. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations[J]. BMC Genet, 2010, 11(1): 94.
- [38] JIANG L, CHEN Y, ZHANG J, et al. Population structure of large yellow croaker (*Larimichthys crocea*) revealed by single nucleotide polymorphisms[J]. Biochem Syst Ecol, 2015, 63: 136-142.
- [39] 李光华, 金方彭, 周睿, 等. 基于 SNP 标记的短须裂腹鱼自然群体遗传多样性分析 [J]. 水生生物学报, 2018, 42(2): 271-276.
- [40] 金方彭, 李光华, 冷云, 等. 澜沧江中上游光唇裂腹鱼四个地理群体遗传多样性分析 [J]. 水生生物学报, 2021, 45(1): 60-68.
- [41] SHAKLEE J B, TAMARU C S, WAPLES R S. Speciation and evolution of marine fishes studied by the electrophoretic analysis of proteins[J]. Pac Sci, 1982, 36(2): 141-157.
- [42] WRIGHT S. Evolution and the genetics of population variability within and among natural populations[M]. Chicago: University of Chicago Press, 1978: 439-459.
- [43] 赵凯, 何舜平, 彭作刚, 等. 青海湖裸鲤的种群结构和线粒体 DNA 变异 [J]. 青海大学学报 (自然科学版), 2006, 24(4): 1-4.
- [44] 张春霖, 陈大庆, 史建全, 等. 青海湖裸鲤繁殖群体遗传多样性的 RAPD 分析 [J]. 水产学报, 2005, 29(3): 307-312.
- [45] 周玲, 金章东, 李福春, 等. 青海湖裸鲤 (湟鱼) 耳石的矿物组成及其 Sr/Ca 对洄游习性的潜在示踪 [J]. 中国科学: 地球科学, 2012, 42(8): 1210-1217.
- [46] 刘伟, 罗仕立, 周其椿, 等. 青海湖裸鲤遗传多样性研究概况 [J]. 农技服务, 2017, 34(17): 86-88.
- [47] 杨海乐, 危起伟. 论水生野生动物的主动保护与被动保护 [J]. 湖泊科学, 2021, 33(1): 1-10.