

DOI: 10.20103/j.stxb.202302220304

邹艳婷, 胡丹心, 吴非霏, 闵星月, 李飞龙, 张远. 珠江流域常见鱼类及大型底栖动物 DNA 条形码空缺分析. 生态学报, 2024, 44(4): 1564-1574.

Zou Y T, Hu D X, Wu F F, Min X Y, Li F L, Zhang Y. DNA barcode gap analysis for common fish and macroinvertebrate species in the Pearl River. Acta Ecologica Sinica, 2024, 44(4): 1564-1574.

珠江流域常见鱼类及大型底栖动物 DNA 条形码空缺分析

邹艳婷¹, 胡丹心², 吴非霏¹, 闵星月³, 李飞龙^{1,*}, 张 远¹

¹ 广东工业大学生态环境与资源学院, 广东省流域水环境治理与水生生态修复重点实验室, 广州 510006

² 广东省广州生态环境监测中心站, 广州 510006

³ 辽宁大学环境学院, 沈阳 110036

摘要: 条形码数据库是开展基于 DNA 的生物监测关键先决条件。为在珠江流域有效开展基于 DNA 的生物监测, 迫切需要了解物种 DNA 条形码的覆盖或空缺状况。整理了珠江流域常见鱼类和大型底栖动物的物种清单, 从 National Center and Biotechnology Information (NCBI) 数据库中检索了物种清单的 DNA 条形码序列, 分析了常见鱼类 (包括线粒体组和 12s rRNA 基因) 和大型底栖动物 (包括线粒体组、COI 和 18s rRNA 基因) 的 DNA 条形码覆盖范围和空缺程度。数据分析表明: (1) 珠江流域共记录了常见鱼类 221 种, 隶属于 2 纲 18 目 51 科和 137 属; 常见大型底栖动物 105 种/属, 隶属于 6 纲 14 目 53 科。(2) 共检索到常见鱼类线粒体组序列 913 条和 12s rRNA 基因序列 962 条, 分别占总物种的 81.45% 和 57.92%; 有 12.67% 的物种没有线粒体组和 12s rRNA 基因序列, 若将条形码阈值设置为至少包含 5 个参考序列, 则空缺度上升至 52.94%; (3) 共检索到常见大型底栖动物线粒体组 65 条序列、COI 基因 26,988 条序列和 18s rRNA 基因 175 条序列, 分别占总种/属数的 29.52%、68.57% 和 37.14%; 有 25.71% 的种/属在线粒体组、COI 和 18s rRNA 基因区域皆无序列收录, 若将条形码阈值设置为至少包含 5 个参考序列, 则空缺度上升至 41.90%。总之, 本研究将为珠江流域开展基于 DNA 的鱼类和大型底栖动物监测提供基础数据支撑, 为完善珠江本土 DNA 条形码数据库提供参考建议。

关键词: 环境 DNA; 线粒体组; COI; 12s rRNA; 18s rRNA

DNA barcode gap analysis for common fish and macroinvertebrate species in the Pearl River

ZOU Yanting¹, HU Danxin², WU Feifei¹, MIN Xingyue³, LI Feilong^{1,*}, ZHANG Yuan¹

¹ Guangdong Provincial Key Laboratory of Water Quality Improvement and Ecological Restoration for Watersheds, School of Ecology, Environment and Resources, Guangdong University of Technology, Guangzhou 510006, China

² Guangzhou Ecological and Environmental Monitoring Center of Guangdong Province, Guangzhou 510006, China

³ School of Environment, Liaoning University, Shenyang 110036, China

Abstract: Barcode reference databases are an important foundation for DNA-based biomonitoring and an indispensable part of the current trending and promising environmental DNA (eDNA) technology. For effective DNA-based biomonitoring in the Pearl River Basin, the coverage or gap status of DNA barcodes of species is urgently needed to be captured. Here, we compiled a detailed historical record and species checklist of common fish and macroinvertebrates in the Pearl River. All available DNA sequences (e.g., the mitochondrial genome, Cytochrome c Oxidase subunit I (COI) gene, 12s rRNA gene,

基金项目: 国家自然科学基金项目 (52239005, 52100216)

收稿日期: 2023-02-22; **网络出版日期:** 2023-11-27

* 通讯作者 Corresponding author. E-mail: fl_li@gdut.edu.cn

and 18s rRNA gene) matching the species names in the checklist of the common fish and macroinvertebrates were downloaded separately in the NCBI GenBank database, and analyzed the coverage of the DNA barcode records of species in the NCBI GenBank database. Our data showed that: (1) a total of 221 common fish species were recorded, including 2 classes, 18 orders, 51 families, and 137 genera, there are 2 endemic genera and 17 endangered rare species, and the common fish mainly included Cypriniformes, Perciformes, Siluriformes, the remaining 15 orders accounted for a combined 24.89%; A total of 6 classes, 14 orders, 53 families, and 105 species (or genera) of common macroinvertebrates in the Pearl River were retrieved, mainly Ephemeroptera, Trichoptera, Plecoptera, the remaining 11 orders accounted for a combined 24.77%. (2) A total of 913 sequences of the mitochondrial genome and 962 sequences of 12s rRNA gene were retrieved, covering 81.45% and 57.92% of the total number of common fish species, respectively; 12.67% of the species had neither mitochondrial group sequences nor 12s rRNA gene sequences. The gap would rise to 52.94% if the barcode thresholds were set to contain at least five reference sequences per taxon. (3) A total of 65 sequences of the mitochondrial genome, 26988 sequences of COI gene and 175 sequences of 18s rRNA gene were retrieved, covering 29.52%, 68.57% and 37.14% of the total number of common macroinvertebrates species (or genera), respectively. 25.71% of the species (or genera) had no sequences included in the mitochondrial group, COI and 18s rRNA gene regions. The gap would rise to 41.90% if the barcode thresholds were set to contain at least five reference sequences per taxon. Overall, the collation of data in this study will provide importantly basic data support for DNA-based monitoring of fish and macrobenthos in the Pearl River Basin, and provide reference suggestions for improving the native DNA barcode database in the Pearl River.

Key Words: eDNA; mitochondrial genome; Cytochrome c Oxidase subunit I; 12s rRNA; 18s rRNA

珠江孕育了丰富的水生生物资源,但随着社会经济的快速发展,航运、渔业捕捞等高强度的人类活动已造成珠江生物多样性急剧下降^[1-3]。大量鱼种从过去的优势物种变为当前的珍稀濒危物种(如花鳗鲡 *Anguilla marmorata*、唇鲮 *Semilabeo notabilis* 等),洄游鱼类种群数量急剧下降,中华鲟(*Acipenser sinensis*)已消失多年^[4]。罗非鱼(*Oreochromis mossambicus*)、麦瑞加拉鲮(*Cirrhinus mrigala*)和胡子鲇(*Clarias fuscus*)等外来物种的入侵进一步加剧了珠江鱼类多样性的下降,严重威胁着珠江水生态安全^[5]。开展快速、高效的生物监测成为珠江流域水生态保护的关键前提,但传统的监测方法费时费力且需要分类鉴定专家^[6-10]。近年来,环境 DNA 技术正引领生物监测进入新时代^[11-15],该技术具有省时省力,高精度的物种识别率等诸多优势^[16-19]。环境 DNA 技术在流域推广应用的先决条件是需要一个完整的本土物种 DNA 条形码数据库^[20-22],若条形码数据库不完整,会使测序得到的 DNA 序列无法进行物种注释^[23-27],进而低估或错估了多样性状况,影响对生态系统中物种的真实分布与丰度的准确认识^[28]。

目前,全球物种 DNA 条形码数据库有 NCBI GenBank、Barcode of Life Data Systems(BOLD)和 EMBL 核酸序列数据库等^[29-31]。其中,NCBI GenBank 和 BOLD 数据库已分别记录超过 1100 万和 2000 万的动物参考序列。加拿大是首个建立了国家物种 DNA 条形码数据库的国家,随后德国、中国、日本等国家也相继开展了本国的物种 DNA 条形码数据库的构建^[32-34]。Weigand 等人^[35]分析了泛欧洲大陆水生物种的 DNA 条形码数据库的覆盖率,发现用于生物监测的条形码数据库空缺在不同分类群和国家之间有明显差异;Vieira 等人^[36]研究了马卡罗尼亚海洋大型底栖动物的 COI 和 18s rRNA 条形码在 BOLD 数据库中的覆盖情况,发现 COI 对不同的分类群覆盖率差异较大;Li 等人^[26]分析了中国河流淡水鱼类、水生昆虫和软体动物的 COI 条形码分别在 NCBI GenBank 和 BOLD 数据库中的覆盖率,指出中国三大水生动物类群远没有一个具有代表性和完整的数据库。上述这些研究均表明 DNA 条形码数据库在不同地理区域或国别、不同分类群以及不同基因区域间存在明显差异。欧美发达地区和受关注的物种 DNA 条形码覆盖率更高,而欠发达地区及其生物多样性热点流域的 DNA 条形码覆盖率缺乏足够了解。因此,为提升流域的水生生物监测能力,将环境 DNA 纳入生物监测业务范畴之前,迫切需要对重点流域的物种 DNA 条形码数据库进行整理和分析^[37-39]。

本研究以珠江流域为研究区域,选取了流域水生态监测中两个重要指示类群:鱼类和大型底栖动物为研究对象^[40-43],分析了珠江流域常见鱼类和大型底栖动物 DNA 条形码覆盖率状况。首先,从全球生物多样性信息基金(GBIF)数据库、已发表的论文和地方鱼类志中检索了珠江常见鱼类和大型底栖动物的物种清单。其次,根据物种清单,从 NCBI GenBank 数据库中检索并下载了常见鱼类(包括线粒体组和 12s rRNA 基因)和大型底栖动物(包括线粒体组、COI 和 18s rRNA 基因)的 DNA 序列。随后,分析了不同基因区域的 DNA 条形码数据库在珠江流域常见鱼类和大型底栖动物中的覆盖率。本研究的核心目的是整理和分析珠江流域常见鱼类及大型底栖动物的物种名录,及其在现有数据库中 DNA 条形码记录现状,并指出不同分类群的条形码覆盖率。本研究收集的 DNA 条形码序列将为在珠江流域开展基于 DNA 的鱼类和大型底栖动物监测提供基础数据支撑,对缺乏 DNA 条形码记录的物种整理将进一步指引着珠江本土 DNA 条形码数据库的完善。

1 数据来源及研究方法

1.1 记录检索及物种清单

为获取珠江流域有明确历史记载的鱼类和大型底栖动物的物种记录,首先在 GBIF 数据库中进行检索,检索区域设置为“China”,以“Actinopterygii”为关键词获取鱼类中辐鳍鱼纲的物种信息;分别以“Mollusca”和“Insecta”为关键词获取大型底栖动物中软体动物和昆虫的物种信息。其次,通过下载的数据表中提供的地理坐标信息截取珠江流域常见鱼类和大型底栖动物物种数据清单。随后,结合研究学者近年来在国内外学术期刊发表的关于珠江流域鱼类和大型底栖动物相关研究、《珠江鱼类志》^[44]、《东江流域水环境与水生态研究》^[45]和《珠江水系鱼类原色图集》^[46]等文献与书籍,进一步完善和补充珠江流域鱼类和大型底栖动物的物种清单。为了对鱼类和大型底栖动物物种的拉丁名进行标准化,包括去掉同义词和检查名称的有效性,将整理的物种清单与 Fishbase、中国生物物种名录 2020 版和维基百科等进行了比较,选择有明确记录的且更常用的名字。最后,编制了珠江流域常见鱼类和大型底栖动物物种清单。

1.2 下载并分析条形码数据库

根据编制的珠江流域常见鱼类和大型底栖动物物种清单,从 NCBI GenBank 数据库中下载所有可用的 DNA 序列,其中,常见鱼类包括线粒体组和 12s rRNA 基因,常见大型底栖动物包括线粒体组、COI 和 18s rRNA 基因。参考已发表文献的要求^[26],仅保留了长度超过 500 bp 且具有明确物种名的 DNA 条形码记录。设定 DNA 条形码最小数量阈值(即 1 条和 5 条),分别计算 DNA 条形码在鱼类和大型底栖动物中的覆盖率(图 1)。

2 结果分析

2.1 珠江常见鱼类和大型底栖动物物种组成

共检索到珠江流域常见鱼类 221 种,隶属于 2 纲 18 目 51 科和 137 属(图 2)。其中,软骨鱼纲(Chondrichthyes)2 种,包含鲛形目(Perciformes)1 目 2 科 2 属(即赤魮 *Dasyatis akabei* 和日本燕魮 *Gymura japonica* Temminck et Schlegel);硬骨鱼纲(Osteichthyes)219 种,隶属于 17 目 49 科 135 属。鲤形目(Cypriniformes)种数最多(97 种),约占珠江鱼类总物种的 43.89%,鲈形目(Perciformes)其次(53 种),约占总物种的 23.89%,鼠鱚目(Gonorhynchiformes)、鲉形目(Scorpaeniformes)及鲿形目(Cyprinodontiformes)占比最少(均为 0.45%)。此外,研究发现了珠江鱼类特有属 2 个,分别为金线鲃属(*Sinocyclocheilus* sp.)和云南鳅属(*Yunnanilus* sp.);珍稀濒危鱼类 17 种,分别为中华鲟(*Acipenser sinensis*),阳宗白鱼(*Anabarilius yangzonensis*),星云白鱼(*Anabarilius andersoni*),唐鱼(*Tanichthys albonubes*),常氏吻孔鲃(*Poropuntius chonglingchungi*),昆明裂腹鱼(*Schizothorax grahami*),抚仙鲤(*Mesocyprius fuxianensis*),云南鲤(*Cyprinus yunnanensis*),翘嘴鲤(*Cyprinus ilishaectomus*),黄唇鱼(*Bahaba taipingensis*),阳宗金线鲃(*Sinensia yangzongensis*),唇鲮(*Semilabeo notabilis*),花鳗鲡(*Anguilla marmorata*)等。

共检索到珠江常见大型底栖动物 105 种/属,隶属于 6 纲 14 目 53 科(图 2)。其中,蜉蝣目

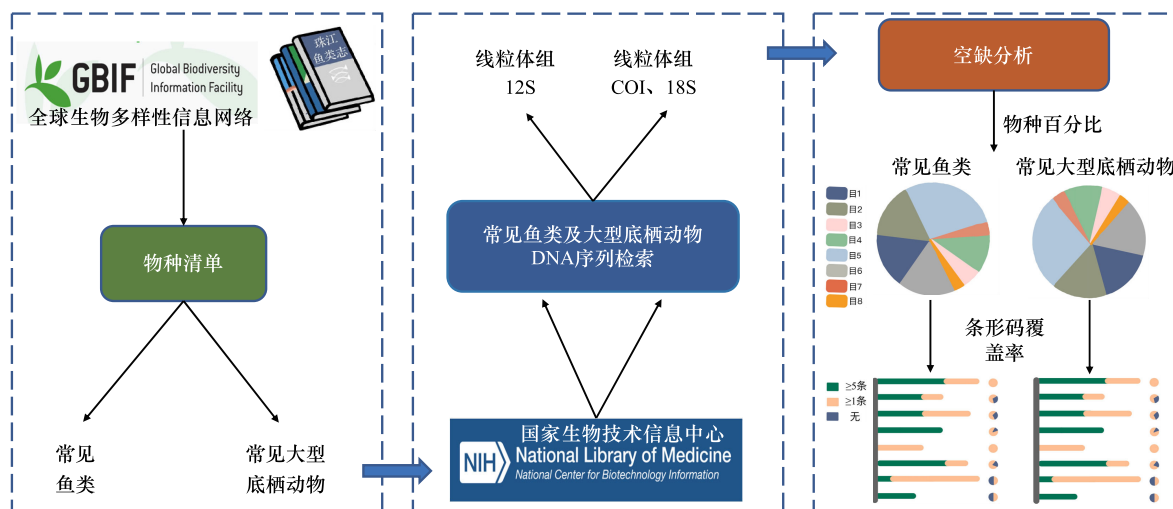


图1 珠江流域常见鱼类和大型底栖动物物种清单和 DNA 序列数据来源及检索方法

Fig.1 Data sources and retrieval methods of species lists and DNA sequences for fish and macroinvertebrates in the Pearl River

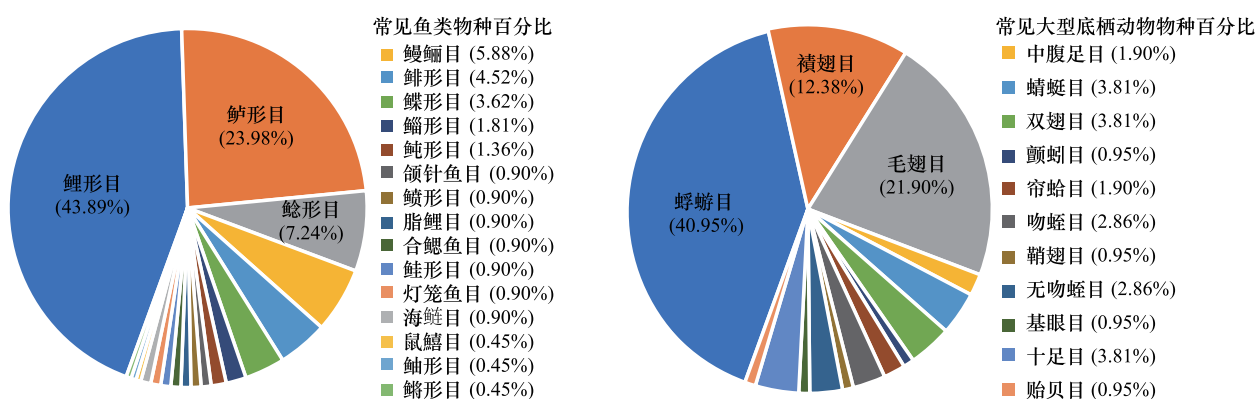


图2 珠江常见鱼类和大型底栖动物的物种组成百分比

Fig.2 Species composition of common fish and macroinvertebrates in the Pearl River

(Ephemeroptera) 43 种/属, 约占总种/属数的 40.95%; 其次是毛翅目 (Trichoptera) 23 种/属, 约占总种/属数的 21.90%; 襁翅目 (Plecoptera) 13 种/属, 约占总种/属数的 12.38%; 其余 11 目共占总种/属数的 24.77%, 其中无吻蛭目 (Arhynchobdellida)、基眼目 (Basommatophora)、十足目 (Decapoda) 以及贻贝目 (Mytiloida) 仅占比 0.95% (图 2)。

2.2 常见鱼类 DNA 条形码数据库

共检索到 913 条线粒体组基因序列, 有 81.45% 的常见鱼类 (共 221 种) 在数据库中至少有 1 条具有代表性的基因序列 (图 3), 其隶属于 18 目 48 科, 占总目数的 100%, 占总科数的 94.12%。不同目间的线粒体组覆盖存在显著差异。其中灯笼鱼目 (Myctophiformes)、海鲢目 (Elopiformes)、鼠鱗目 (Gonorhynchiformes)、脂鲤目 (Characiformes)、合鳃鱼目 (Synbranchiformes)、鲑形目 (Salmoniformes)、鲃形目 (Tetraodontiformes)、鲉形目 (Scorpaeniformes) 和 鲈形目 (Cyprinodontiformes) 的鱼类线粒体组基因序列覆盖率为 100%; 鲱形目 (Siluriformes) 有 93.75% 的物种存在线粒体组序列, 鲱形目 (Actinopterygii) 也有 90% 的物种有一个具有代表性的线粒体组序列。鲤形目 (Cypriniformes) 物种数最多 (97 种), 其中 83 个物种有线粒体组序列, 线粒体组基因序列覆盖率为 85.57%。颌针鱼目 (Beloniformes)、鳊形目 (Myliobatiformes) 和 鲃形目 (Mugiliformes) 的线粒体

组基因序列覆盖率最低,仅占该目物种种数的 50%;57.92%的常见鱼类(总物种 221 种)在 NCBI GenBank 数据库中至少有 1 条具有代表性的 12s rRNA 基因序列(图 3),隶属于 18 目 46 科,占总目数的 100%,占总科数的 90.20%。灯笼鱼目(Myctophiformes)、海鲢目(Elopiformes)、合鳃鱼目(Synbranchiformes)、鼠鲮目(Gonorhynchiformes)、鲑形目(Salmoniformes)、鲈形目(Tetraodontiformes)、鲉形目(Scorpaeniformes)和鲈形目(Cyprinodontiformes)的 12s rRNA 基因序列覆盖率均为 100%,其次是鲈形目(Perciformes)(覆盖率 86.79%)、鲈形目(Actinopterygii)(覆盖率 80%)和鳗鲡目(Anguilliformes)(覆盖率 69.23%)等。鲤形目(Cypriniformes) 12s rRNA 基因序列数最多,但覆盖率仅 38.14%。鲶形目(Siluriformes)的 12s rRNA 基因序列覆盖率最低,只有 25%的物种有 12s rRNA 基因序列收录(图 3)。

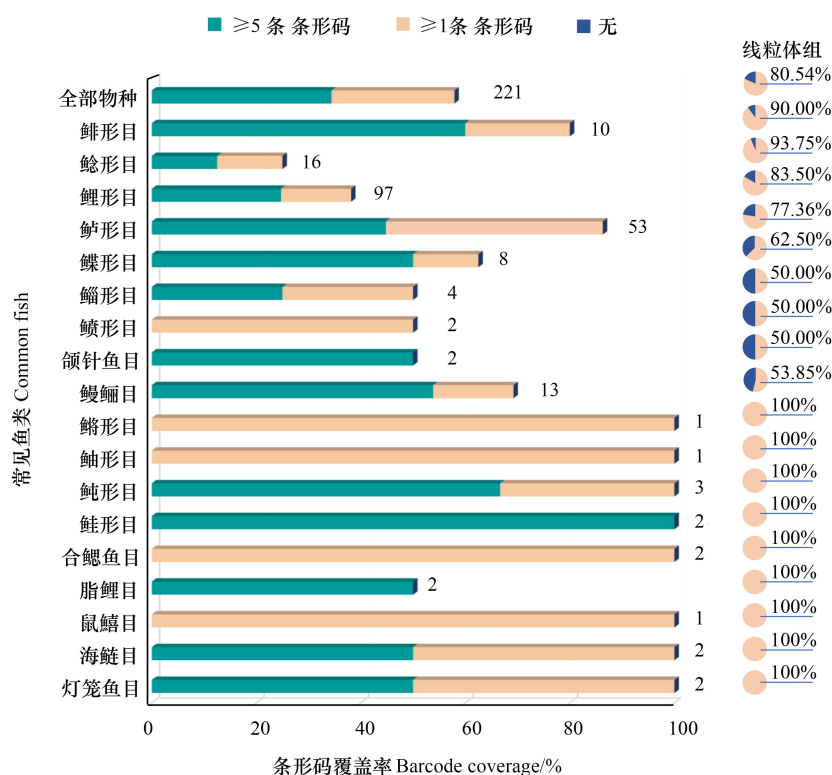


图 3 常见鱼类线粒体组和 12s rRNA 基因序列的空缺分析,条形图上的数字表示物种数

Fig.3 Gap analysis of the mitochondrial genome and 12s rRNA gene of common fish. The numbers on the bars are the species number

如果考虑到包含至少 5 个条形码序列的,总物种的覆盖率下降到 23.53%,即共 52 个物种有 5 条或更多 12s rRNA 基因序列,隶属于 15 目(图 3)。其中,合鳃鱼目(Synbranchiformes)、鼠鲮目(Gonorhynchiformes)、鲉形目(Scorpaeniformes)和鲈形目(Cyprinodontiformes)中所有常见鱼类的 12s rRNA 基因序列数皆 ≥ 5 。此外,数据显示有 27 个物种可检索的 12s rRNA 基因序列只有 1 条记录,隶属于 9 目,物种数占总种数的 12.11%,包括蜥海鲢(*Elops saurus*)、花鲈(*Clupanodon thrissa*)、黄泽小沙丁鱼(*Sardinella lemuru*)、圆吻海鲢(*Nematalosa nasus*)、团头鲂(*Megalobrama amblycephala*)、花鲈(*Hemibarbus maculatus*)、东方墨头鱼(*Garra orientalis*)、革胡子鲶(*Clarias gariepinus*)、间下鱗鱼(*Hyporhamphus intermedius*)、圆鳞斑鲆(*Pseudorhombus levisquamis*)、弓斑多纪鲷(*Takifugu ocellatus*)等。

虽然在同时考虑线粒体组和 12s rRNA 基因区域时,常见鱼类 DNA 条形码覆盖率达到 87.33%,但仍有 28 个物种没有线粒体组和 12s rRNA 基因序列收录,隶属于 9 目。主要包括有:星云白鱼(*Anabarilius andersoni*)、日本燕魮(*Gymnura japonica*)、乌耳鳗鲡(*Anguilla nigricans*)、中华须鳗(*Cirrhimuraena chinensis*)、尖吻蛇鳗(*Ophichthus apicalis*)、细身光唇鱼(*Acrossocheilus elongates*)、花鲈(*Tephriactes sinensi*)、斑点薄鳅(*Leptobotia punctatus*)、黑鳉(*Sarcocheilichthys nigripinnis nigripinnis*)等。

2.3 常见大型底栖动物 DNA 条形码数据库

共检索到常见大型底栖动物线粒体组基因序列 65 条,来自于 31 种/属,仅占总种/属数(105 种/属)的 29.52%。其中,基眼目(Basommatophora)的线粒体组基因序列覆盖率为 100%,其次是鞘翅目(Coleoptera),覆盖率为 50%。蜉蝣目(Ephemeroptera)中 18 种/属收录有线粒体组基因序列,占该目种/属数的 41.86%,且有 4 个种/属的线粒体组收录基因序列数 ≥ 5 。毛翅目(Trichoptera)有 6 种/属收录到线粒体组基因序列,占该目种/属数的 26.09%,仅有 1 个种/属线粒体组收录基因序列数 ≥ 5 。襁翅目(Plecoptera)有 2 种/属收录到线粒体组基因序列,占该目种/属数的 15.38%,但它们都仅有 1 条基因序列收录。无吻蛭目(Arhynchobdellida)、吻蛭目(Rhynchobdellida)、颤蚓目(Tubificida)、贻贝目(Mytiloida)、中腹足目(Mesogastropoda)以及十足目(Decapoda)暂无线粒体组基因序列收录。此外,共 17 个种/属可检索到的线粒体组基因序列只有 1 条记录,隶属于 5 目,占总种/属数的 16.19%,占可检索序列种/属的 54.84%,包含绢蜉(*Ephemera serica*)、美丽高翔蜉(*Epeorus melli*)、扁蜉属(*Heptagenia* sp.)、等蜉属(*Isonychia* sp.)、红纹蜉属(*Rhoenanthus* sp.)、广西河花蜉(*Potamanthus kwangsiensis*)、大别山越南蜉(*Vietnamella dabieshanensis*)、倍叉属(*Amphinemura* sp.)、缘脉多距石蛾属(*Plectrocnemia* sp.)、原石蛾属(*Rhyacophila* sp.)、叶春蛭属(*Ictinogomphus* sp.)等。

共检索到 26,988 条 COI 基因序列,隶属于 72 种/属,占常见大型底栖动物总种/属数的 68.57%。其中,鞘翅目(Coleoptera)、双翅目(Diptera)、无吻蛭目(Arhynchobdellida)、吻蛭目(Rhynchobdellida)、颤蚓目(Tubificida)、贻贝目(Mytiloida)、基眼目(Basommatophora)、中腹足目(Mesogastropoda)以及十足目(Decapoda)的 COI 基因序列覆盖率都是 100%。其次是毛翅目(Trichoptera),有 21 个种/属能在 NCBI GenBank 数据库中找到 COI 基因序列,占该目种/属数的 91.30%,且其中 19 个种/属有 5 个或者更多的 COI 基因序列,占该目种/属数的 82.61%。此外,毛翅目(Trichoptera)收录 COI 基因序列 5459 条,其大部分种/属皆收录有丰富的 COI 基因序列,如:小石蛾属(*Hydroptila* sp.)收录 COI 基因序列 1127 条;合脉石蛾属(*Cheumatopsyche* sp.)收录 COI 基因序列 850 条;*Hydropsyche* sp. 收录 COI 基因序列 766 条;缺叉石蛾属(*Chimarra* sp.)收录 COI 基因序列 444 条等。

种/属数量最多的蜉蝣目(Ephemeroptera)中有 23 个种/属在数据库中可查到至少 1 条 COI 基因序列,覆盖率为 53.49%,其中有 18 个种/属有 5 个或者更多的条形码,覆盖率为 41.86%,收录 COI 基因序列共有 2000 条,部分种/属 COI 基因序列收录丰富,如小蜉属(*Ephemerella* sp.)收录 COI 基因序列 625 条;高翔蜉属(*Epeorus* sp.)收录 COI 基因序列 322 条;梧州蜉(*Ephemera wuchowensis*)收录 COI 基因序列 216 条等。襁翅目(Plecoptera)有 5 个种/属能在 NCBI GenBank 数据库中检索到 COI 基因序列,仅占该目种/属数的 38.46%,收录 COI 基因序列 376 条,且科与科之间的基因序列覆盖率相差较大。例如,叉襁科(Nemouridae)的 3 个种/属都能在 NCBI GenBank 数据库中找到 COI 基因序列,且多于 5 条;襁科(Perlidae)仅有 2 个种/属有记录且少于 2 条;襁翅目(Plecoptera)中其余 4 个科都无法检索到 COI 基因序列。

NCBI GenBank 数据库中共收录了 175 条常见大型底栖动物的 18s rRNA 基因序列,占总种/属数的 34.21%(共 105 种/属)(图 4)。其中,鞘翅目(Coleoptera)、无吻蛭目(Arhynchobdellida)、十足目(Decapoda)、颤蚓目(Tubificida)以及基眼目(Basommatophora)中种/属皆可查到 18s rRNA 基因序列,其次是中腹足目(Mesogastropoda)和双翅目(Diptera),均有 75%的种/属可查 18s rRNA 基因序列。蜉蝣目(Ephemeroptera)有 12 个种/属在数据库中可查到至少 1 条 18s rRNA 基因序列,占该目种/属数的 26.67%。毛翅目(Trichoptera)有 9 个种/属在数据库中可查到至少 1 条 18s rRNA 基因序列,覆盖率为 31.03%,但是毛翅目(Trichoptera)中种/属的可查 18s rRNA 基因序列数都小于 5。襁翅目(Plecoptera)只有 1 个种/属在数据库中可查到 18s rRNA 基因序列,覆盖率仅为 7.14%。吻蛭目(Rhynchobdellida)和贻贝目(Mytiloida)暂无 18s rRNA 基因序列的收录。

常见大型底栖动物中有 9 个种/属的 18s rRNA 基因序列 ≥ 5 ,隶属于 6 目,占总种/属数的 7.89%,分别为四节蜉属(*Baetis* sp.)、弯握蜉属(*Drunella* sp.)、小蜉属(*Ephemerella* sp.)、纯蛭属(*Paragnetina* sp.)、狭溪泥甲

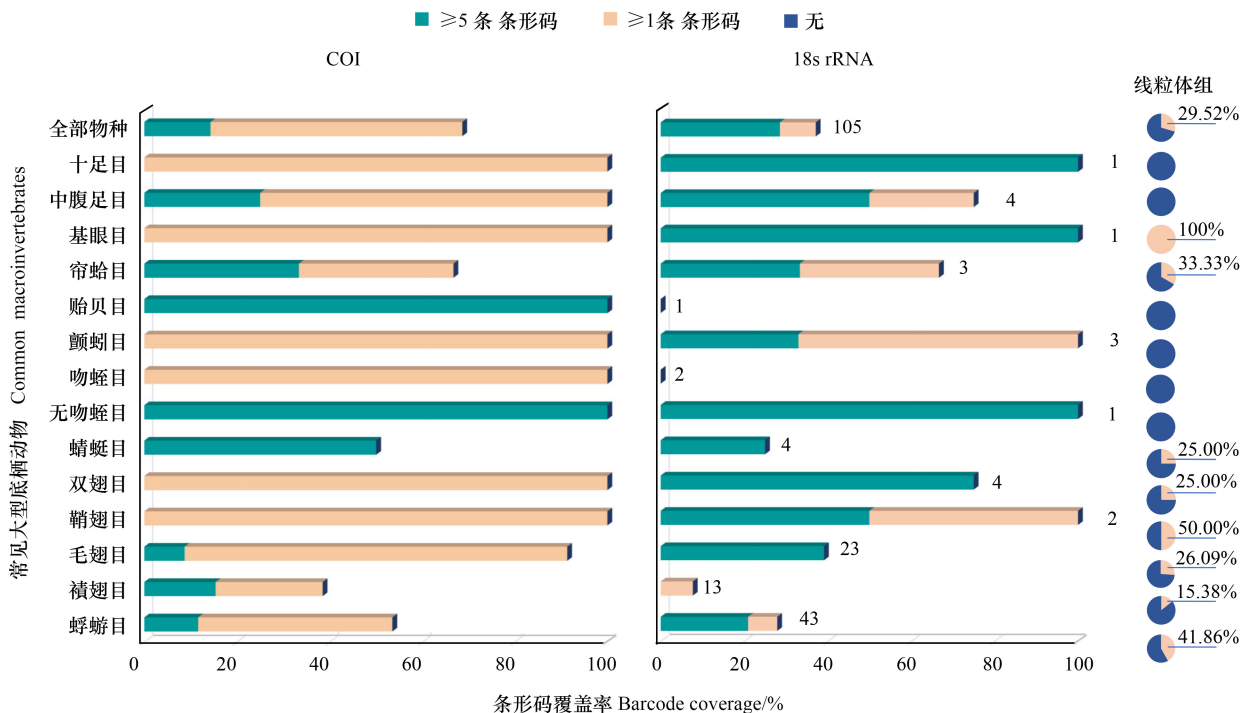


图4 常见大型底栖动物线粒体组、COI 以及 18s rRNA 基因序列的空缺分析

Fig.4 Gap analysis of the mitochondrial genome, COI gene and 18s rRNA gene of common macroinvertebrates

属 (*Stenelmis* sp.)、水丝蚓属 (*Limnodrilus* sp.)、*Nais* sp.、伞菌属 (*Corbicula* sp.)、环棱螺属 (*Bellamya* sp.)。13 个种/属可检索的 18S rRNA 基因序列只有 1 条记录,隶属于 7 目,占总种/属数的 11.40%,包括似动蜉属 (*Cinygmia* sp.)、高翔蜉属 (*Epeorus* sp.)、红纹蜉属 (*Rhoenanthus* sp.)、*Oecetis* sp.、缘脉多距石蛾属 (*Plectrocnemia* sp.)、*Hydrobiosidae* sp.、叶春蜉属 (*Ictinogomphus* sp.)、巴蜉属 (*Barbronia* sp.)、*Radix* sp. 和 *Palaemon* sp.等。

数据显示有 27 个种/属无线粒体组、COI 以及 18s rRNA 基因序列收录记录,占总种/属数的 25.71%,隶属于 5 目,包括假二翅蜉属 (*Pseudocloeon* sp.)、宝加带肋蜉 (*Cincticostella boja*)、红天角蜉 (*Uracanthella rufa*)、徐氏蜉 (*Ephemera hsui*)、宽基蜉属 (*Choroterpes* sp.)、小裳蜉属 (*Leptophlebia* sp.)、吉氏柔裳蜉 (*Habrophlebiodes gilliesi*)、中华细蜉 (*Caenis sinensis*)、钩襀属 (*Kamimuria* sp.)、中叉襀属 (*Mesonemura* sp.)等。

3 讨论

本研究通过对 GBIF 数据库、发表或出版的文献和书籍检索,整理出珠江流域常见鱼类 221 种和大型底栖动物 105 种/属,分析了这些物种在 NCBI GenBank 数据库中的 DNA 条形码的覆盖率。数据显示,即使综合考虑本研究中的各条形码区域,珠江常见鱼类和大型底栖动物中分别有 12.67% 和 25.71% 的物种仍然缺少条形码参考序列。如果将条形码收录的阈值设置为 ≥ 5 条,那么条形码空缺度将上升到 52.94% 的常见鱼类和 41.90% 的常见大型底栖动物。因对种/属水平的识别有较高的分辨率,目前线粒体组、12s rRNA、18s rRNA 和 COI 基因片段已被广泛用于鱼类和大型底栖动物的监测^[47-50]。本研究针对上述基因区域 DNA 条形码数据的整理,可以相对全面地评估珠江流域常见鱼类和大型底栖动物的 DNA 条形码现状。然而,数据显示的 DNA 条形码空缺程度意味着珠江流域尚缺乏一个完整的条形码数据库,将极大的限制基于 DNA 的生物监测在珠江流域的应用。

珠江常见鱼类基因序列整体覆盖率较好,87.33% 的物种收录有至少 1 条基因序列,但大多物种的序列数

都较少(<5 条)。仅 23.53%和 18.55%的鱼类有不少于 5 条的 12s rRNA 基因序列和线粒体组基因序列收录。某一物种在数据库中如果仅收录一条或少数序列将影响 DNA 序列的物种注释准确率。一方面,因为缺乏有效监管,数据库中单一序列本身可能就存在错误,进而增加物种注释的错配概率;另一方面,单一或少数的序列通常覆盖的地理区域过于狭窄,无法全面地代表物种自身遗传进化和基因多样性,这就导致对于不同地区的同一物种中的遗传差异无法预估^[51-52]。因此,条形码数据库需要对序列数少于 5 条的物种进行补充测序研究,在物种分布的不同地理区域中至少包含多条 DNA 条形码序列,以满足对生物多样性精准监测需求。

珠江常见大型底栖动物 DNA 条形码数据目前主要为 COI 基因序列,线粒体组及 18s rRNA 研究非常少。而 COI 基因序列数虽然相当可观,但其在不同目中的序列数差异非常大,双翅目(Diptera)和毛翅目(Trichoptera)收录的序列数就占了数据库中 80%以上的数据记录,而其他 12 个目中仍有许多种/属序列数不超过 5 条,甚至没有序列收录,这表明珠江常见大型底栖动物中大部分物种只有一个基因区域的条形码或没有序列收录。众所周知,大型底栖动物是生态系统中多样性最为丰富的类群之一,其中的物种进化关系极为复杂,软体动物和寡毛类种间遗传距离较大且难以判断,仅依靠单一片段进行识别易造成分类错误^[53]。如,有研究发现 *Baetis* sp.和 *Oligochaeta* 的物种 COI 序列存在超过 10%的遗传差异,而一些昆虫的遗传差异仅为 1%—3%^[54]。因此,条形码数据库还需要收录多个基因的条形码来克服由于基于单一区域进行识别而导致的分类错位^[55-56]。今后研究中应进行更广泛的采样,对只有 COI 基因序列收录的物种补充线粒体组和 18s rRNA 的测序,对目前还未收录有序列的物种应更加关注,尤其是已知是生态中重要组成部分的物种^[57]。大型底栖动物中覆盖率差异较大的分类群也应重点关注,补充其中覆盖率较低的物种的条形码序列,建立一个更为全面的条形码数据库^[58-59]。

针对物种 DNA 条形码空缺的困境,建立一个全面的本土物种条形码数据库应需分别从技术和管理层面开展。在技术层面上,第三代测序技术、基因捕获等技术的最新进展,不仅有效降低了测序成本和提高测序通量^[60],而且便携、快速的应用场景使得在野外即可完成对物种 DNA 序列的测序^[61-62]。单分子测序无需 PCR、长的序列读长(> 1kb)和无 GC 偏好等特点,解决了由于扩增不良或偏好、碱基重复复杂导致的测序质量差等问题^[63]。新技术的掌握与应用将为填补物种条形码空白提供快速有效的方法与手段。在管理层面上,除关注每个物种的多个序列之外,对 DNA 序列的仔细汇编、验证和注释也是建立可靠数据库的基础^[64]。目前公共数据库上传数据的门槛较低,导致数据库中的数据质量参差不齐,对于其中可能存在的标记错误、序列冲突、分类冲突等问题无法预估^[35,64]。因此,DNA 条形码数据库管理中应提高质量标准,对分类分配进行交叉验证或标记可疑的条形码,同时还须考虑这些序列的来源、地理信息等,避免由于本地条形码与非本地条形码的成对遗传距离差异导致的不匹配,确保物种由准确的 DNA 序列进行表示^[65]。通过对不同地区或不同物种进行分开管理,建立更多本地化数据库或针对单一物种的序列库,进而改善条形码数据库的质量控制和管理工作流程^[66-67]。

总的来说,本研究不仅整理了珠江流域常见鱼类和大型底栖动物的物种名录,分析了这些物种 DNA 条形码的覆盖状况,而且明确指出了哪些物种在现有的数据库中仍缺乏 DNA 条形码记录,强调了填补条形码数据库空白的紧迫性,分别从技术和管理层面给出了相关的建议和策略,这将为在珠江流域开展基于 DNA 的生物监测提供基础数据支持。

4 结论

(1) 根据 GBIF 数据库及文献资料,珠江共记录了常见鱼类 221 种和常见大型底栖动物 105 种/属。常见鱼类中有 2 个特有属,17 种濒危珍稀物种;常见大型底栖动物中大多为蜉蝣目、毛翅目。

(2) 在 NCBI GenBank 数据库中共检索到珠江流域常见鱼类的线粒体组基因组序列 913 条和 12s rRNA 基因序列 962 条,分别占常见鱼类总物种数的 81.45%和 57.92%,其中 12.67%的物种既没有线粒体组序列也没有 12s rRNA 基因序列收录,且仅 23.53%和 18.55%的鱼类有不少于 5 条的 12s rRNA 基因序列和线粒体组

基因序列收录。常见鱼类的条形码数量不足将影响 DNA 序列的物种注释准确率,后续应补充条形码序列收录。

(3)在 NCBI GenBank 数据库中共检索到常见大型底栖动物线粒体组序列 65 条、COI 基因序列 26,988 条和 18s rRNA 基因序列 175 条,分别占常见大型底栖动物总种/属数的 29.52%、68.57%和 37.14%,其中有 25.71%的种/属线粒体组、COI 和 18s rRNA 基因序列皆无收录。大型底栖动物的序列主要为单一的 COI 基因序列,应补充线粒体组和 18s rRNA 的测序。

参考文献(References):

- [1] Chen W T, Li C, Li X H, Li J, Li Y F. Unraveling the drifting larval fish community in a large spawning ground in the middle Pearl River using DNA barcoding. *Animals: an Open Access Journal from MDPI*, 2022, 12(19): 2555.
- [2] Kuang T X, Chen W J, Huang S H, Liu L, Zhou L. Environmental drivers of the functional structure of fish communities in the Pearl River Estuary. *Estuarine, Coastal and Shelf Science*, 2021, 263: 107625.
- [3] Xia Y G. Ecological and economic impacts of exotic fish species on fisheries in the Pearl River Basin. *Management of Biological Invasions*, 2019, 10(1): 127-138.
- [4] 帅方敏, 李新辉, 何安尤, 刘乾甫, 张迎秋, 武智, 朱书礼. 珠江水系广西江段鱼类多样性空间分布特征. *水生生物学报*, 2020, 44(4): 819-828.
- [5] 高天扬, 谢迪, 彭宁东, 张少平, 李潮, 罗锦桢, 王春晖, 赵俊. 北江鱼类群落结构多样性及其演替趋势. *水生态学杂志*, 2018, 39(4): 54-62.
- [6] Shuai F, Lek S, Li X, Zhao T. Biological invasions undermine the functional diversity of fish community in a large subtropical river. *Biological Invasions*, 2018, 20(10): 2981-2996.
- [7] Li C, Wang J J, Chen J Q, Schneider K, Veettil R K, Elmer K R, Zhao J. Native bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix* populations in the Pearl River are threatened by Yangtze River introductions as revealed by mitochondrial DNA. *Journal of Fish Biology*, 2020, 96(3): 651-662.
- [8] van der Plas F. Biodiversity and ecosystem functioning in naturally assembled communities. *Biological Reviews*, 2019: brv.12499.
- [9] Shuai F M, Li J. Nile Tilapia (*Oreochromis niloticus* Linnaeus, 1758) invasion caused trophic structure disruptions of fish communities in the South China River-pearl River. *Biology*, 2022, 11(11): 1665.
- [10] Guo D L, Zhou L, Wang G P, Lai H, Bi S, Chen X L, Zhao X P, Liu S, Luo Y, Li G F. Use of artificial structures to enhance fish diversity in the Youjiang River, a dammed river of the Pearl River in China. *Ecology and Evolution*, 2020, 10(23): 13439-13450.
- [11] McGee K M, Robinson C V, Hajibabaei M. Gaps in DNA-based biomonitoring across the globe. *Frontiers in Ecology and Evolution*, 2019, 7: 337.
- [12] Elías-Gutiérrez M, Hubert N, Collins R A, Andrade-Sossa C. Aquatic organisms research with DNA barcodes. *Diversity*, 2021, 13(7): 306.
- [13] Miller S E, Hausmann A, Hallwachs W, Janzen D H. Advancing taxonomy and bioinventories with DNA barcodes. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 2016, 371(1702): 20150339.
- [14] 李飞龙, 杨江华, 杨雅楠, 张效伟. 环境 DNA 宏条形码监测水生生态系统变化与健康状态. *中国环境监测*, 2018, 34(6): 37-46.
- [15] Yang C H, Wu K C, Chuang L Y, Chang H W. DeepBarcoding: deep learning for species classification using DNA barcoding. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022, 19(4): 2158-2165.
- [16] Pawlowski J, Apothéoz-Perret-Gentil L, Altermatt F. Environmental DNA: what's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, 2020, 29(22): 4258-4264.
- [17] Pawlowski J, Bonin A, Boyer F, Cordier T, Taberlet P. Environmental DNA for biomonitoring. *Molecular Ecology*, 2021, 30(13): 2931-2936.
- [18] Leray M, Knowlton N, Ho S L, Nguyen B N, Machida R J. GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(45): 22651-22656.
- [19] Gostel M R, Kress W J. The expanding role of DNA barcodes: indispensable tools for ecology, evolution, and conservation. *Diversity*, 2022, 14(3): 213.
- [20] Paz G, Rinkevich B. Gap analysis of DNA barcoding in ERMS reference libraries for ascidians and cnidarians. *Environmental Sciences Europe*, 2021, 33(1): 4.
- [21] Kress W J, García-Robledo C, Uriarte M, Erickson D L. DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution*, 2015, 30(1): 25-35.
- [22] Fontes J T, Vieira P E, Ekrem T, Soares P, Costa F O. BAGS: an automated Barcode, Audit & Grade System for DNA barcode reference libraries.

- Molecular Ecology Resources, 2021, 21(2): 573-583.
- [23] Leite B R, Vieira P E, Teixeira M A L, Lobo-Arteaga J, Hollatz C, Borges L M S, Duarte S, Troncoso J S, Costa F O. Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. *Regional Studies in Marine Science*, 2020, 36: 101307.
- [24] Tzafesta E, Saccomanno B, Zangaro F, Vadrucchi M R, Specchia V, Pinna M. DNA barcode gap analysis for multiple marker genes for phytoplankton species biodiversity in Mediterranean aquatic ecosystems. *Biology*, 2022, 11(9): 1277.
- [25] Porter T M, Hajibabaei M. Over 2.5 million COI sequences in GenBank and growing. *PLoS One*, 2018, 13(9): e0200177.
- [26] Li F L, Zhang Y, Altermatt F, Zhang X W, Cai Y P, Yang Z F. Gap analysis for DNA-based biomonitoring of aquatic ecosystems in China. *Ecological Indicators*, 2022, 137: 108732.
- [27] Huemer P, Mutanen M. An incomplete European barcode library has a strong impact on the identification success of Lepidoptera from Greece. *Diversity*, 2022, 14(2): 118.
- [28] Lin X L, Mo L D, Bu W J, Wang X H. The first comprehensive DNA barcode reference library of Chinese Tanytarsus (Diptera: Chironomidae) for environmental DNA metabarcoding. *Diversity and Distributions*, 2021, 27(10): 1932-1941.
- [29] Meiklejohn K A, Damaso N, Robertson J M. Assessment of BOLD and GenBank-Their accuracy and reliability for the identification of biological materials. *PLoS One*, 2019, 14(6): e0217084.
- [30] Duarte S, Leite B, Feio M, Costa F, Filipe A. Integration of DNA-based approaches in aquatic ecological assessment using benthic macroinvertebrates. *Water*, 2021, 13(3): 331.
- [31] Thakur M, Bateman A, Brooksbank C, Freeberg M, Harrison M, Hartley M, Keane T, Kleywegt G, Leach A, Levchenko M, Morgan S, McDonagh E M, Orchard S, Papatheodorou I, Velankar S, Vizcaino J A, Witham R, Zdrazil B, McEntyre J. EMBL's European bioinformatics institute (EMBL-EBI) in 2022. *Nucleic Acids Research*, 2023, 51(D1): D9-D17.
- [32] Trivedi S, Aloufi A A, Ansari A A, Ghosh S K. Role of DNA barcoding in marine biodiversity assessment and conservation; an update. *Saudi Journal of Biological Sciences*, 2016, 23(2): 161-171.
- [33] Hawlitschek O, Morinière J, Dunz A, Franzen M, Rödger D, Glaw F, Haszprunar G. Comprehensive DNA barcoding of the herpetofauna of Germany. *Molecular Ecology Resources*, 2016, 16(1): 242-253.
- [34] Tanizawa Y, Fujisawa T, Kodama Y, Kosuge T, Mashima J, Tanjo T, Nakamura Y. DNA Data Bank of Japan (DDBJ) update report 2022. *Nucleic Acids Research*, 2023, 51(D1): D101-D105.
- [35] Weigand H, Beermann A J, Čiampor F, Costa F O, Csabai Z, Duarte S, Geiger M F, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand A M, Willassen E, Wyler S A, Bouchez A, Borja A, Čiamporová-Zatovičová Z, Ferreira S, Dijkstra K D B, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn B B, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher J N, Mamos T, Paz G, Pešić V, Pfannkuchen D M, Pfannkuchen M A, Price B W, Rinkevich B, Teixeira M A L, Várbró G, Ekrem T. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *The Science of the Total Environment*, 2019, 678: 499-524.
- [36] Vieira P E, Lavrador A S, Parente M I, Parretti P, Costa A C, Costa F O, Duarte S. Gaps in DNA sequence libraries for Macaronesian marine macroinvertebrates imply decades till completion and robust monitoring. *Diversity and Distributions*, 2021, 27(10): 2003-2015.
- [37] Specchia V, Tzafesta E, Marini G, Scarcella S, D'Attis S, Pinna M. Gap analysis for DNA barcode reference libraries for aquatic macroinvertebrate species in the Apulia region (southeast of Italy). *Journal of Marine Science and Engineering*, 2020, 8(7): 538.
- [38] Phillips J D, Gillis D J, Hanner R H. Lack of statistical rigor in DNA barcoding likely invalidates the presence of a true species' barcode gap. *Frontiers in Ecology and Evolution*, 2022, 10: 859099.
- [39] Jin S, Kim K Y, Kim M S, Park C. An assessment of the taxonomic reliability of DNA barcode sequences in publicly available databases. *ALGAE*, 2020, 35(3): 293-301.
- [40] Múrria C, Somma S, Wangenstein O S, Arnedo M A, Prat N. Towards an Iberian DNA barcode reference library of freshwater macroinvertebrates and fishes. *Limnetica*, 2020, 39(1): 73-92.
- [41] 王萌, 苑艺, 于海燕, 贾显乐, 黄浩然, 陈月仙, 金小伟, 林晓龙, 王备新. 中国淡水大型底栖无脊椎动物条形码数据库构建. *中国环境监测*, 2022, 38(1): 36-44.
- [42] 金珂, 张丽娟, 张伟, 张翔, 陈桥, 杨江华, 张咏, 张效伟. 基于环境 DNA 宏条形码的太湖流域底栖动物监测与生态健康评价. *中国环境监测*, 2022, 38(1): 175-188.
- [43] Jin L N, Yu J, Yuan X Q, Du X S. Fish classification using DNA barcode sequences through deep learning method. *Symmetry*, 2021, 13(9): 1599.
- [44] 郑慈英. 珠江鱼类志. 北京: 科学出版社, 1989.

- [45] 杨扬, 王赛, 崔永德. 东江流域水环境与水生态研究. 北京: 科学出版社, 2020.
- [46] 李新辉, 陈方灿, 梁沛文. 珠江水系鱼类原色图集: 广东段. 北京: 科学出版社, 2018.
- [47] Andújar C, Arribas P, Gray C, Bruce C, Woodward G, Yu D W, Vogler A P. Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, 2018, 27(1): 146-166.
- [48] Elyasigorji Z, Izadpanah M, Hadi F, Zare M. Mitochondrial genes as strong molecular markers for species identification. *The Nucleus*, 2023, 66(1): 81-93.
- [49] 王晨, 陶孟, 李爱民, 施鹏, 杨江华, 王志浩, 张效伟. 基于环境 DNA 宏条形码技术的秦淮河生物多样性研究. *生态学报*, 2022, 42(2): 611-624.
- [50] Zangaro F, Saccomanno B, Tzafesta E, Bozzeda F, Specchia V, Pinna M. Current limitations and future prospects of detection and biomonitoring of NIS in the Mediterranean Sea through environmental DNA. *NeoBiota*, 2021, 70: 151-165.
- [51] Shen Y J, Hubert N, Huang Y, Wang X Z, Gan X N, Peng Z G, He S P. DNA barcoding the ichthyofauna of the Yangtze River: insights from the molecular inventory of a mega-diverse temperate fauna. *Molecular Ecology Resources*, 2019, 19(5): 1278-1291.
- [52] Tsuji S, Shibata N, Inui R, Nakao R, Akamatsu Y, Watanabe K. Environmental DNA phylogeography: successful reconstruction of phylogeographic patterns of multiple fish species from cups of water. *Molecular Ecology Resources*, 2023, 23(5): 1050-1065.
- [53] Curry C J, Gibson J F, Shokralla S, Hajibabaei M, Baird D J. Identifying North American freshwater invertebrates using DNA barcodes: are existing COI sequence libraries fit for purpose? *Freshwater Science*, 2018, 37(1): 178-189.
- [54] Vivien R, Wyler S, Lafont M, Pawlowski J. Molecular barcoding of aquatic oligochaetes: implications for biomonitoring. *PLoS One*, 2015, 10(4): e0125485.
- [55] Martins F M S, Porto M, Feio M J, Egeter B, Bonin A, Serra S R Q, Taberlet P, Beja P. Modelling technical and biological biases in macroinvertebrate community assessment from bulk preservative using multiple metabarcoding markers. *Molecular Ecology*, 2021, 30(13): 3221-3238.
- [56] Antil S, Abraham J S, Sripoorna S, Maurya S, Dagar J, Makhija S, Bhagat P, Gupta R, Sood U, Lal R, Toteja R. DNA barcoding, an effective tool for species identification: a review. *Molecular Biology Reports*, 2023, 50(1): 761-775.
- [57] Jazdzewska A M, Tandberg A H S, Horton T, Brix S. Global gap-analysis of amphipod barcode library. *PeerJ*, 2021, 9: e12352.
- [58] Hestetun J T, Bye-Ingebrigtsen E, Nilsson R H, Glover A G, Johansen P O, Dahlgren T G. Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Marine Biodiversity*, 2020, 50(5): 70.
- [59] Pinna M, Saccomanno B, Marini G, Zangaro F, Kabayeva A, Khalaj M, Shaimardan L, D'Attis S, Tzafesta E, Specchia V. Testing the influence of incomplete DNA barcode libraries on ecological status assessment of Mediterranean transitional waters. *Biology*, 2021, 10(11): 1092.
- [60] Lyons E, Sheridan P, Tremmel G, Miyano S, Sugano S. Large-scale DNA barcode library generation for biomolecule identification in high-throughput screens. *Scientific Reports*, 2017, 7: 13899.
- [61] Srivathsan A, Baloglu B, Wang W, Tan W X, Bertrand D, Ng A H Q, Boey E J H, Koh J J Y, Nagarajan N, Meier R. A MinION™-based pipeline for fast and cost-effective DNA barcoding. *Molecular Ecology Resources*, 2018, 18: 1035-1049.
- [62] Truelove N K, Andruszkiewicz E A, Block B A. A rapid environmental DNA method for detecting white sharks in the open ocean. *Methods in Ecology and Evolution*, 2019, 10(8): 1128-1135.
- [63] Roberts R J, Carneiro M O, Schatz M C. The advantages of SMRT sequencing. *Genome Biology*, 2013, 14(6): 1-4.
- [64] Keck F, Couton M, Altermatt F. Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. *Molecular Ecology Resources*, 2023, 23(4): 742-755.
- [65] Blackman R C, Walser J C, Rüber L, Brantschen J, Villalba S, Brodersen J, Seehausen O, Altermatt F. General principles for assignments of communities from eDNA: open versus closed taxonomic databases. *Environmental DNA*, 2023, 5(2): 326-342.
- [66] Csabai Z, Čiamporová-Zatovičová Z, Boda P, Čiampor F. 50%, not great, not terrible: pan-European gap-analysis shows the real status of the DNA barcode reference libraries in two aquatic invertebrate groups and points the way ahead. *Science of the Total Environment*, 2023, 863: 160922.
- [67] Rahman M M, Norén M, Mollah A R, Kullander S O. Building a DNA barcode library for the freshwater fishes of Bangladesh. *Scientific Reports*, 2019, 9: 9382.