

基于线粒体 *COI* 和 *18S rRNA* 基因构建舟山海域浮游动物 DNA 宏条形码数据库

冼华琳, 汤昌盛, 张杨阳, 张晓林, 廖 智, 周钰琴, 谢舒烨, 严小军
浙江海洋大学 海洋科学与技术学院, 浙江 舟山 316022

摘要: 为构建舟山海域浮游动物本地条形码数据库, 从而为该海域浮游动物的种类鉴定、系统进化关系及种群遗传学研究提供科学依据。基于过去 14 年的文献报道, 编制了收录 23 目 64 科 96 属 155 种浮游动物的舟山海域浮游动物物种名录, 通过从 NCBI GenBank 参考数据库获取相关数据, 共收集浮游动物线粒体 *COI* 基因序列 195 条和核糖体 *18S rRNA* 基因序列 134 条, 初步建立了舟山海域浮游动物条形码数据库。对最优势类群桡足类进行了种间和种内遗传距离的计算分析。结果显示: 桡足类 *COI* 基因序列种间平均遗传距离为 0.313 6, 是种内平均遗传距离的 34.09 倍; *18S rRNA* 基因序列的种间平均遗传距离为 0.262 8, 是种内平均遗传距离的 26.82 倍。*COI* 基因序列的种间和种内遗传距离存在明显的条形码间隔, 而 *18S rRNA* 基因序列的种间和种内遗传距离存在重叠, 没有明显的条形码间隔。基于 *COI* 和 *18S rRNA* 基因序列构建的系统发育树表明, 2 种条形码基因在目和科水平上均能有效区分不同的浮游动物类群。舟山海域浮游动物 DNA 宏条形码数据库的建立, 有利于浮游动物种类鉴定及多样性评估, 并为该海域海洋生态环境状况评价提供科学依据。

关键词: 浮游动物; *COI* 基因; *18S rRNA* 基因; DNA 条形码数据库; 舟山海域

中图分类号: Q 953

文献标志码: A

开放科学(资源服务)标识码(OSID):



Construction of DNA metabarcoding database of zooplankton in Zhoushan sea area based on mitochondrial cytochrome *COI* and *18S rRNA* gene

XIAN Hualin, TANG Changsheng, ZHANG Yangyang, ZHANG Xiaolin, LIAO Zhi, ZHOU Yuqin, XIE Shuyue, YAN Xiaojun

Marine Science and Technology College, Zhejiang Ocean University, Zhoushan 316022, China

Abstract: In order to create an indigenous DNA metabarcoding database for zooplankton in the Zhoushan sea area and support the research on species identification, phylogenetic relationships, and genetic analysis of marine zooplankton, we compiled a list of zooplankton species in the Zhoushan sea area based on data collected over the past 14 years. We downloaded a total of 195 mitochondrial cytochrome oxidase subunit I (*COI*) barcode sequences and 134 small ribosomal subunit *18S rRNA* barcode sequences from the NCBI GenBank database. Then we calculated the intra- and interspecific genetic distances of copepod species. The results show that the average interspecific genetic distance of the *COI* gene was 0.313 6, which was 34.09 times as high as the average distance within each copepod species. Likewise, the average interspecific genetic distance for the *18S rRNA* gene was 0.262 8, which was 26.82 times as high as the average intraspecific distance. The DNA barcode of the *COI* gene exhibited a clear barco-

收稿日期: 2025-01-11; 修回日期: 2025-04-09

基金项目: 国家自然科学基金国际(地区)合作研究与交流项目(42020104009); 国家自然科学基金青年科学基金项目(32200083); 浙江省基础公益研究计划项目(LTGS23C010001); 浙江海洋大学人才引进基金项目(JX6311104324)

作者简介: 冼华琳(1999—), 男, 硕士研究生, 研究方向为浮游动物生态学。E-mail: 15766529280@163.com

通信作者: 汤昌盛(1989—), 男, 讲师, 博士, 研究方向为海洋生物学。E-mail: tangchsh@zjou.edu.cn

ding gap, while there were overlaps between the intra-and interspecific genetic distances based on the 18S *rRNA* gene. The phylogenetic tree constructed using *COI* and 18S *rRNA* genes demonstrates that both genes can distinguish different zooplankton taxa at the order and family levels effectively. This DNA barcoding database will facilitate species identification and biodiversity assessment of zooplankton, providing valuable references for evaluating the marine environment in the Zhoushan sea area.

Keywords: Zooplankton; *COI* gene; 18S *rRNA* gene; DNA barcoding database; Zhoushan sea area

舟山海域位于浙江省东北部,地处杭州湾外缘的东海。流经该海域的台湾暖流、长江冲淡水 and 浙闽沿岸流等水团、上升流等带来了丰富的营养物质^[1],为浮游生物的迅速繁殖创造了良好的条件,进一步为鱼、虾等海洋生物提供了充足的饵料,舟山渔场因势形成。近年来,舟山市水产养殖业和旅游业等产业快速发展,随之而来的环境污染、过度捕捞、养殖侵占等给海洋生态系统带来巨大压力,舟山渔场资源也面临衰竭^[2]。保护和恢复海洋生态系统,改善海洋生态环境质量,对于实现海洋的可持续发展显得尤为重要。

浮游动物是水生生态系统的重要组成部分,既作为鱼类等更高营养级的食物来源,也能通过摄食作用调控浮游植物群落的生物量、物种组成等^[3],在食物链的能量流动和物质循环上起着承上启下的作用^[4]。此外,浮游动物具有生命周期短、代谢活动强等特性^[5],对环境变化非常敏感,常作为指示种评价海域水质状况,如北冰洋和大西洋中桡足类丰度的变化可揭示两大洋中水团相互作用的变化^[6]。因此,了解研究区域中浮游动物的种类组成、丰度、多样性等群落特征可为海域的生态系统健康评估、生物多样性保护、管理和决策等提供科学依据。

传统的浮游动物物种鉴定方法主要为形态学鉴定,依赖于浮游动物的体节数、附肢数量和形态、尾叉、个体大小、复眼、心脏等特征作为判定依据。但形态学鉴定方法对鉴定者的经验要求很高,且无法同时鉴定多个样品,需要耗费大量的时间^[7];而浮游动物中也存在很多表型相似的隐存种^[8],浮游幼体尤其是桡足类幼体也很难通过形态学方法鉴定到种水平^[9],无法准确评估浮游动物的物种多样性。分子技术的发展实现了利用 DNA 宏条形码技术实现浮游动物的快速鉴定。DNA 宏条形码技术依托高通量测序平台,通过扩增样品中特定的标准化短基因片段^[10],如线粒体细胞色素 C 氧化酶亚基 I (Cytochrome oxidase subunit I, *COI*) 基因、18S 核糖体 RNA (18S ribosomal RNA, 18S

rRNA)、16S 核糖体 RNA (16S ribosomal RNA, 16S *rRNA*) 等基因的可变区得到操作分类单元 (Operational taxonomic units, OUT) 或扩增子序列变体 (Amplicon sequence variants, ASV) 后^[11],与参考数据库中的参考序列进行比对以实现物种注释。由于目前使用的 GenBank、BOLD 等公共参考数据库中存在很多冗余和重复的序列^[12],这些冗余的数据往往来自于不同研究者对同一物种测序结果的重复提交,或者由于技术误差、样本污染等问题导致序列质量不高,影响了物种鉴定的准确性。经验证,浮游动物中小长腹剑水蚤 (*Oithona nana*) 在 GenBank 数据中 10 个长度为 438 bp 的 *COI* 基因序列均为重复序列 (序列号分别为: KU982956.1、KU982955.1、KU982954.1、KU982953.1、KU982952.1、KU982951.1、KU982950.1、KU982949.1、KU982948.1、KU982947.1)。建立区域性本地条形码数据库对特定地区的物种鉴定具有重要意义,其优势主要体现在以下方面:首先是减少冗余,本地条形码数据库可以消除公共数据库中重复或质量不高的序列,提高数据的质量;其次,基于特定区域采集的本地物种数据构建的数据库,能够更精确地匹配目标物种,减少物种鉴定过程中假阳性和假阴性的发生率^[13]。此外,完善的本地 DNA 条形码数据库可为区域生物多样性监测、生态系统评估及物种保护策略制定提供科学依据。

目前,全球已建立了国际海洋浮游动物普查计划 (Census of Marine Zooplankton, CMarz)、覆盖日本北海道、东京湾等海域的日本生物条形码倡议 (Japanese Barcode of Life Initiative, JBOLI) 及覆盖澳大利亚海域的海洋生物多样性中心 (Marine Biodiversity Hub, MBH) 等多个 DNA 条形码数据库,但尚未建立针对舟山海域浮游动物的本地 DNA 条形码数据库,也缺乏对浮游动物各类群系统发育关系分析的报道。因此,本研究根据已有的舟山海域浮游动物群落调查文献编制了舟山海域浮游动物物种名录,并基于 *COI* 和 18S *rRNA* 基因构建本地浮游动物 DNA 宏条形码数据库,评估这 2 种 DNA

条形码基因在舟山海域浮游动物鉴定中的适用性以及浮游动物的系统发育关系, 为后续利用 DNA 宏条形码技术对舟山海域浮游动物的物种鉴定及多样性研究提供科学依据。

1 材料与方法

1.1 数据库构建

整理归纳过去 14 年舟山海域浮游动物资源调查文献^[14-22], 编制舟山海域浮游动物名录, 物种的有效学名以世界海洋物种名录 (World Register of Marine Species) (<https://www.marinespecies.org/>) 为准。根据舟山海域浮游动物物种名录, 在 GenBank 数据库 (<https://www.ncbi.nlm.nih.gov/genbank/>) 上分别下载浮游动物的 *COI* 及 *18S rRNA* 基因序列。为确保选择的序列与舟山海域浮游动物区域特征相匹配, 下载标准是优先下载标注浮游动物采样地点在中国的序列。在 Python 3.8 软件上运行 Biopython 1.85 对下载的 FASTA 格式的序列进行质量控制, 去除重复序列, 以及含有较多简并碱基如 Y 和 K 等的低质量序列。对于有多条序列信息的物种, 使用 SeaView 5.1 软件对该物种全部的序列进行比对, 选取相似性程度最高的 2 条序列, 生成共有序列 (Consensus sequences) 作为该物种的条形码基因代表序列^[23]; 对仅有 1 条序列的物种, 则以该序列作为该物种的代表序列。

1.2 数据分析

在 TBtools-II 软件上将各浮游动物 DNA 条形码基因代表序列合并分析, 进行 MUSCLE 多重序列比对, 并用 TrimAL 2.0 软件将比对后的序列进行剪切以获得一致序列^[24]。在 MEGA 11 上分析序列的核苷酸组成、密码子组成、平均碱基转换/颠换值、保守位点、简约信息位点和变异位点等^[25-26], 并使用 DnaSP 5.0 软件计算序列的单倍型数 (H)、单倍型多样性指数 (H_d)、核苷酸多样性指数 (P_i) 及多态位点数等^[27-28]。浮游动物包含桡足类、毛颚类、浮游多毛类等多个类群, 其中, 桡足类为优势类群^[29]。因为 DNA 条形码基因在不同浮游动物类群间的变异速率存在较大差异^[30], 且目前尚缺乏对桡足类种内、种间遗传差异的报道, 因此本研究仅计算桡足类 *COI* 基因序列及 *18S rRNA* 基因序列的种间及种内遗传距离。本研究选择 Kimura-2-parameter (K2P) 模型用于遗传距离的计算, 该模型

用最广泛的模型, 考虑了碱基转换和颠换速率不同的情况^[31], 尤其适合近缘物种的分析^[32]。使用 MEGA 11 软件基于最大似然法 (Maximum likelihood, ML) 构建分子系统发育树, 并进行 1000 次自展检验 (Bootstraps) 以检验各分支的置信度。生成的 ML 系统发育树在 iTOL 软件 (<http://itol.embl.de/>) 上进行美化 and 注释^[33]。

为更好地评估 *COI* 和 *18S rRNA* 基因在浮游动物物种界定上的可行性及局限性, 采用 ABGD (Automatic Barcode Gap Discovery) 软件 (<https://bioinfo.mnhn.fr/abi/public/abgd/>) 对浮游动物的 *COI* 及 *18S rRNA* 基因序列进行分类单元的划分, 相对间隙宽度值取 $X=0.5$, 其余参数全部默认, 最后比较 ABGD 方法与 ML 系统发育树的物种划分结果。

2 结果

2.1 条形码数据库概况

根据过去 14 年舟山海域浮游动物的种类资源调查报道, 共统计浮游动物 155 种, 隶属于 23 目 64 科 96 属 (图 1, 附录 A, 详见 <http://dx.doi.org/10.12131/20250014> 的资源附件)。其中, 哲水蚤目种类数最多, 共 64 种, 占比 41.3%; 其次为剑水蚤目, 共 17 种, 占比 11.00%; 属于浮游软体动物的有: 管水母目、软水母目、花裸鳃目、硬水母目、筐水母目、淡水水母目、球栉水母目及瓜水母目, 共 23 种, 占比 14.84%。在 GenBank 数据库中下载了 109 种浮游动物的 *COI* 基因序列 195 条及 91 种浮游动物的 *18S rRNA* 基因序列 135 条, 分别覆盖了统计的舟山海域浮游动物中 70.32% 和 58.71% 的种类, 其中有 76 种浮游动物在 GenBank 数据库上有 *COI* 及 *18S rRNA* 2 种基因的 DNA 条形码信息 (图 2)。此外, 在目水平上, 各物种的条形码覆盖率存在差异 (图 3)。其中, *COI* 基因序列在硬水母目、翼足目、钩足目、筐水母目、淡水水母目、软水母目、磷虾目、瓜水母目及花裸鳃目上的覆盖率均达 100%, 而在十足目、端足目和海樽目上的覆盖率则仅为 33.33%、25.00% 和 0, 未超过 50.00% (图 3)。*18S rRNA* 基因序列在硬水母目、钩足目、筐水母目、淡水水母目和瓜水母目的条形码覆盖率为 100%, 但在糠虾目和壮肢目的覆盖率均为 0。

2.2 序列特征

91 种浮游动物的 *18S rRNA* 基因序列经比对,

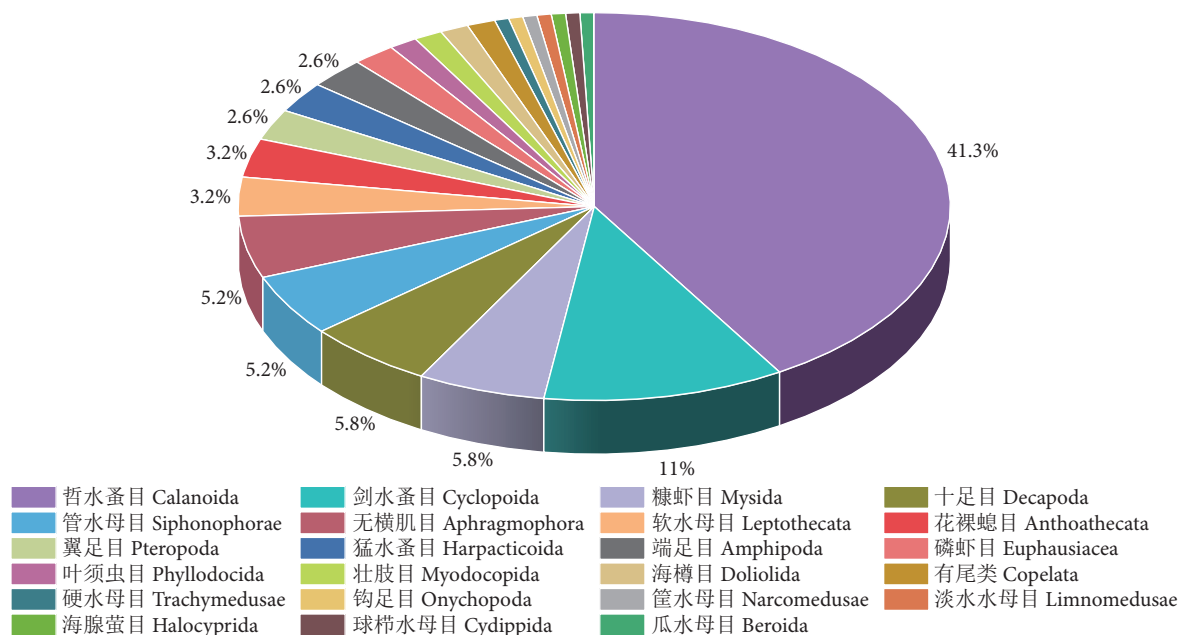


图1 舟山海域浮游动物种类组成

Fig. 1 Zooplankton species composition in Zhoushan sea area

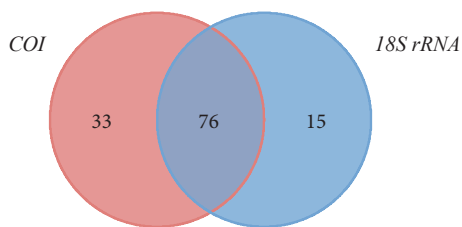


图2 舟山海域浮游动物中具有 DNA 条形码序列 (COI 或 18S rRNA 基因) 的物种数量对比维恩图

Fig. 2 Comparison of zooplankton species numbers with DNA barcode sequences (COI and/or 18S rRNA genes) in Zhoushan sea area by Venn diagram

获得长度为 1 413 bp 的一致序列, 其中可变位点 1 330 个 (94.12%)、简约信息位点 992 个 (70.21%)、保守位点 76 个 (5.38%), 平均碱基转换/颠换值为 1.2。T、C、A、G 的平均含量分别为 25.5%、20.9%、26.8%、26.8%, A+T 碱基占比 52.3%, G+C 碱基占比 47.7%, 具有 A+T 偏倚性。对 63 种桡足类 18S rRNA 基因序列进行比对, 获得一段 1 467 bp 的一致序列, 其中包括 1 248 个可变位点 (85.07%)、721 个简约信息位点 (49.15%)、208 个保守位点 (14.18%), T、C、A、G 的平均含量分别为 25.2%、21.4%、26.7%、26.6%, A+T 的含量 (51.9%) 同样高于 G+C 的含量 (48.0%)。密码子第一位中, A-1 平均含量最高, 占 27.7%; 密码子第二位及第三位则分别是 G-2 和 A-3, 分别占 27.5% 和 27.4%, 存在 A 偏倚性。平均碱基转换/颠换值为 1.1。在 91 种浮游动物和其中 63 种桡足类的 18S rRNA 基

因序列中均只检测出 1 个单倍型, 核苷酸多样性指数分别为 0.214 60 和 0.193 86, 表现出很高的遗传均质性^[34]。

109 种浮游动物的 COI 基因序列经比对, 保留一致序列 831 bp, 其中可变位点 715 个 (86.04%)、简约信息位点 558 个 (67.15%)、保守位点 108 个 (13.00%), T、C、A、G 的平均含量分别为 36.9%、17.0%、25.9%、20.2%, 平均碱基交换/颠换值为 0.8, A+T 的含量为 62.8%, 高于 G+C 的含量 (37.2%), 含量比为 1.69, 具 A+T 偏倚性。单倍型多样性指数和核苷酸多样性指数分别为 0.896 和 0.104 90, 共检测到 46 个单倍型和 35 个多态位点。68 种桡足类的 COI 基因序列获得一致序列 633 bp, 其中可变位点 486 个 (76.78%)、简约信息位点 356 个 (56.24%)、保守位点 147 个 (23.22%), T、C、A、G 的平均含量则分别为 36.5%、17.2%、25.8%、20.6%, A+T 和 G+C 的含量分别为 62.3% 和 37.8%, 同样表现为 A+T 偏倚性。在密码子前三位中, G-1、T-2 和 T-3 的平均含量最高, 分别为 30.5%、44.9% 和 39.3%, 存在 T 偏倚性。平均碱基转换/颠换值为 1.0。单倍型多样性指数和核苷酸多样性指数分别为 0.990 和 0.098 28, 检测出 56 个单倍型和 49 个多态性位点。

2.3 种内及种间遗传距离

基于 K2P 遗传距离模型分别计算桡足类的种内及种间遗传距离 (表 1)。结果表明, 基于 COI 基

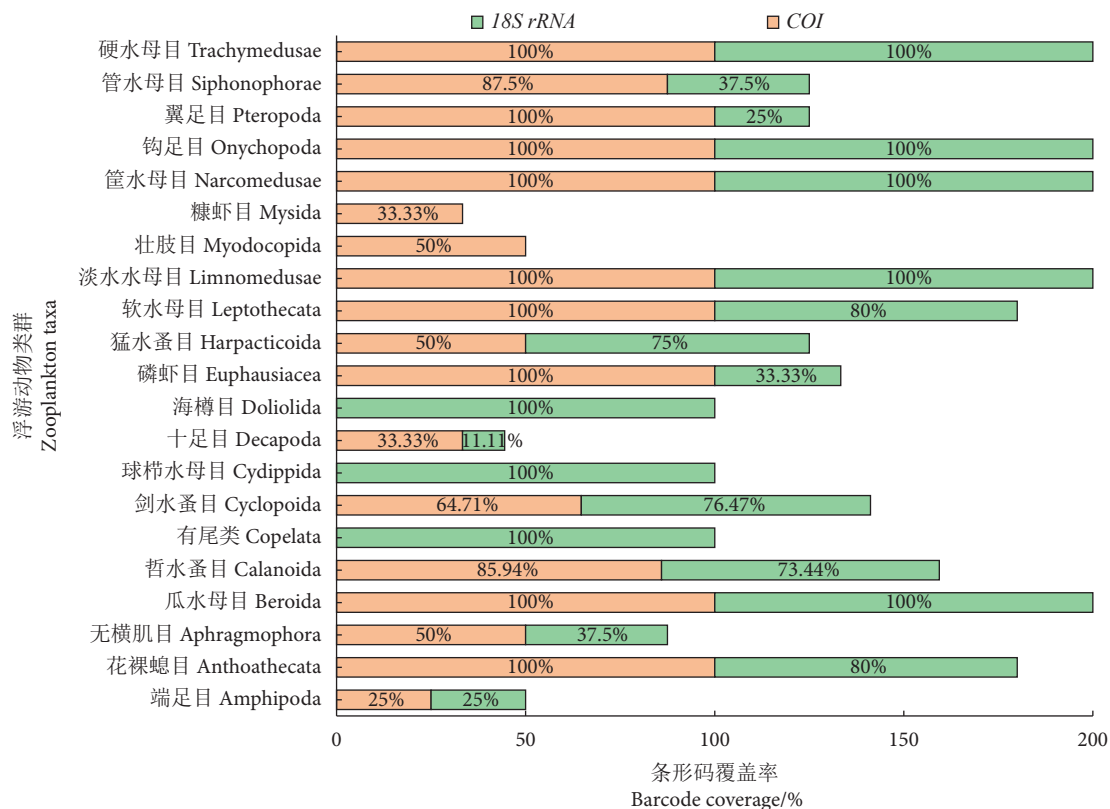


图3 舟山海域浮游动物各类群 DNA 条形码覆盖率对比

Fig. 3 Comparison of DNA barcode coverage among different zooplankton taxa in Zhoushan sea area

因序列, 68 种桡足类的种间平均遗传距离为 0.313 6, 平滑真刺水蚤 (*Euchaeta plana*) 和尖额真猛水蚤 (*Euterpina acutifrons*) 的种间遗传距离最大 (0.994 6), 2 物种分别属于哲水蚤和猛水蚤; 海洋真刺水蚤 (*E. marina*) 和黄角光水蚤 (*Lucicutia flavicornis*) 的种间遗传距离最小 (0.020 1), 两者均为哲水蚤。种

内遗传距离为 0~0.082 0, 平均为 0.009 2, 针刺拟哲水蚤 (*Paracalanus aculeatus*) 的种内遗传距离最大 (0.082 0)。种间平均遗传距离是种内的 34.09 倍, 能形成明显的条形码间隔 (图 4-a), 表明使用 COI 基因作为 DNA 条形码基因能有效区分桡足类物种。

表1 基于 COI 和 18S rRNA 基因序列的桡足类种内与种间遗传距离

Table 1 Intraspecific and interspecific distances of copepods based on COI and 18S rRNA gene sequences

遗传距离 Genetic distance	COI 基因序列 COI gene sequence			18S rRNA 基因序列 18S rRNA gene sequence		
	最小值 Minimum	最大值 Maximum	平均值 Mean	最小值 Minimum	最大值 Maximum	平均值 Mean
种内 Intraspecies	0	0.082 0	0.009 2	0	0.088 2	0.009 8
种间 Interspecies	0.020 1	0.994 6	0.313 6	0.003 1	0.988 7	0.262 8

基于 18S rRNA 基因序列, 63 种桡足类种间平均遗传距离为 0.262 8, 达氏波水蚤 (*Cosmocalanus darwinii*) 和丹氏纺锤水蚤 (*Acartia danae*) 的种间遗传距离最大, 达 0.988 7, 分别隶属于哲水蚤科和纺锤水蚤科; 细长腹剑水蚤 (*Oithona attenuate*) 和小长腹剑水蚤 (*O. nana*) 间的遗传距离最小 (0.003 1), 均属长腹剑水蚤属。种内平均遗传距离为 0.009 8,

最小值为 0, 尖额真猛水蚤的种内遗传距离最大 (0.088 2); 种间平均遗传距离是种内平均遗传距离的 26.82 倍, 满足 Hebert 等^[35] 提出的作为 DNA 条形码分子标记的标记基因其种间遗传距离应为种内遗传距离 10 倍以上的要求, 但基于 18S rRNA 基因序列的遗传距离频数分布柱状图没有形成明显的条形码间隔 (图 4-b), 可能难以对桡足类近缘种进行

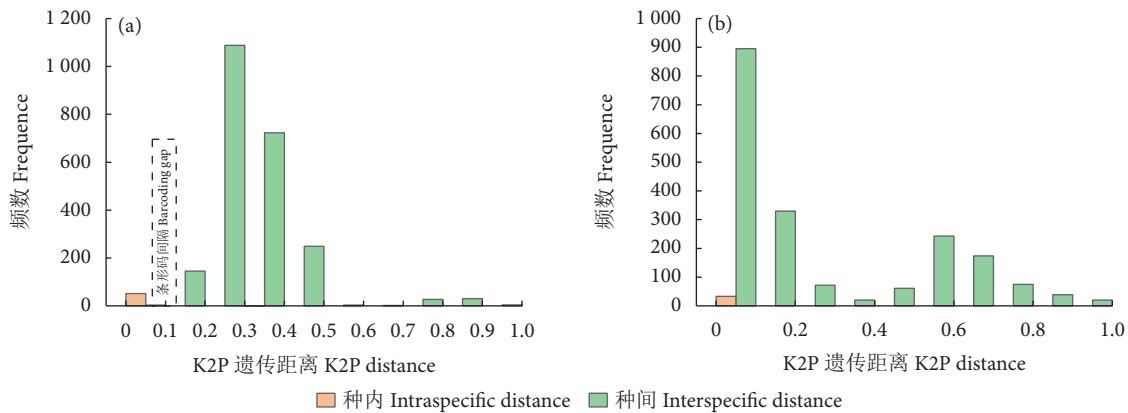


图4 基于 COI (a) 和 18S rRNA (b) 基因序列的桡足类种内和种间 K2P 遗传距离模型的频率分布

Fig. 4 Frequency distributions of intra- and interspecific genetic distances (K2P model) based on COI (a) and 18S rRNA (b) gene sequences of copepods

有效区分。

2.4 系统发育关系

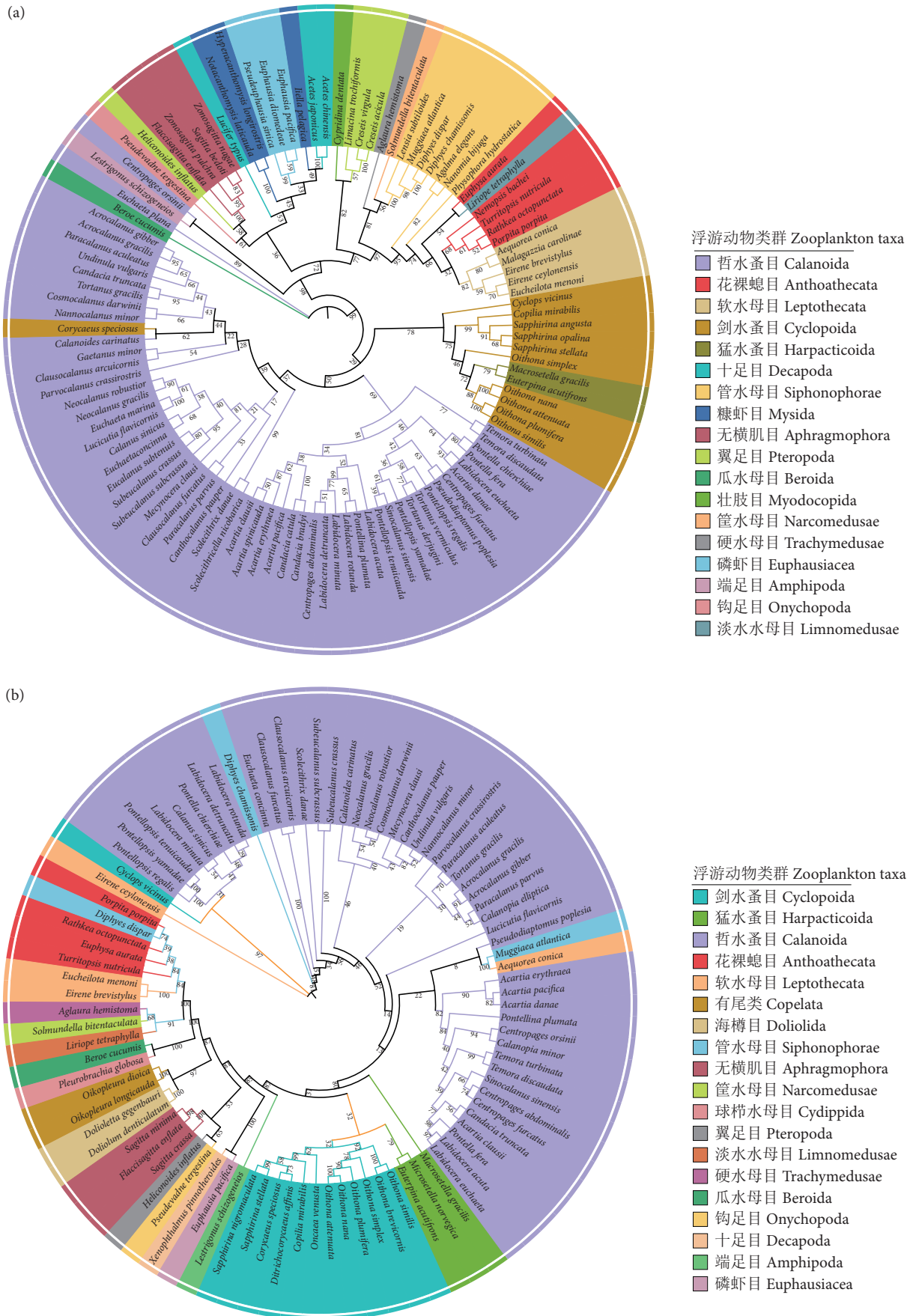
基于 COI 基因序列构建的浮游动物 ML 系统发育树如图 5-a 所示。在目阶元上, 哲水蚤目、剑水蚤目、磷虾目、猛水蚤目、软水母目、管水母目、十足目、无横肌目和花裸螳目的物种均能聚为目的单系支, 涵盖 91.74% 的物种, 表明 COI 基因序列在目阶元上对各浮游动物类群具有较高的分辨率。而仅包含一个物种的壮肢目和瓜水母目则单独成支。同时, 花裸螳目的耳状囊水母 (*Euphysa aurata*) 和淡水水母目的四叶小舌水母 (*Liriope tetraphylla*) 聚为一支, 哲水蚤目的奥氏胸刺水蚤 (*Centropages orsinii*) 与端足目的裂颚蚤戎 (*Lestrigonus schizogeneios*) 聚为一支, 与传统分类存在分歧。

基于 18S rRNA 基因序列构建的浮游动物 ML 系统发育树如图 5-b 所示。在目阶元上, 72.53% 的物种均能与同一目的物种聚成目的单系支。其中, 哲水蚤目、剑水蚤目、猛水蚤目、软水母目和海樽目均能聚为目的单系支。此外, 管水母目的五角水母 (*Muggiaea atlantica*) 和软水母目的锥状多管水母 (*Aequorea conica*) 聚为一支, 瓜水母目的瓜水母 (*Beroe cucumis*) 和球带水母目的球形侧腕水母 (*Pleurobrachia globosa*) 聚为一支, 与传统分类存在差异。结果表明, 18S rRNA 和 COI 基因均能在目阶元上有效区分不同的浮游动物类群, 但 18S rRNA 基因的分辨率要低于 COI 基因。

为进一步评估 COI 和 18S rRNA 基因在科阶元上区分桡足类的能力, 基于 2 种基因序列分别构建 ML 系统发育树。基于 COI 基因序列的 ML 系统发育树表明 (图 6-a), 包括宽水蚤科、哲水蚤科、

长腹剑水蚤科、叶水蚤科、拟哲水蚤科、角水蚤科及平头水蚤科在内, 58.82% 的物种均能聚成科的单系支, 但也存在胸刺水蚤科的奥氏胸刺水蚤、叉胸刺水蚤 (*C. furcatus*) 与其他科的物种聚为一支的情况, 表明 COI 基因未能将胸刺水蚤科的物种与其他科的物种进行有效区分。而基于 18S rRNA 基因序列构建的 ML 系统发育树表明 (图 6-b), 包括拟哲水蚤科、纺锤水蚤科、长腹剑水蚤科、哲水蚤科、叶水蚤科、大眼水蚤科、角水蚤科、真哲水蚤科及宽水蚤科在内 66.67% 的物种均能聚为科的单系支, 而仅有 1 个种的光水蚤科和隆水蚤科则单独为一支, 表明 18S rRNA 基因在科阶元上能对桡足类物种进行有效的鉴定。

通过 ABGD 软件以 0.001 至 0.1 为先验值 p 区间分别对 109 条浮游动物的 COI 基因序列和 91 条 18S rRNA 基因序列, 以及其中的 68 条桡足类 COI 基因序列和 63 条桡足类 18S rRNA 基因序列进行分组, 划分分类单元, 结果如图 7 和图 8 所示。109 条浮游动物 COI 基因序列被划分为 81 个分类单元 (图 7-a), 其中, 隆哲水蚤属、歪水蚤属、刺哲水蚤属、拟哲水蚤属和波水蚤属的物种聚为 1 个分类单元; 水母类的软水母目、淡水水母目、花裸螳目、管水母目的物种则聚为另一个分类单元, 与 ML 系统发育树的分类结果相似。91 条浮游动物的 18S rRNA 基因序列被划分为 41 个分类单元 (图 7-b), 其中, 筒角水蚤属的瘦尾筒角水蚤 (*Pontellopsis tenuicauda*) 和钝筒角水蚤 (*P. yamadae*) 聚为一个分类单元; 住囊虫属的异体位囊虫 (*Oikopleura dioica*) 和长尾住囊虫 (*O. longicauda*) 聚为一个分类单元; 水母类的淡水水母目、硬水母

图5 基于 *COI* (a) 和 *18S rRNA* (b) 基因序列构建的浮游动物 ML 系统发育树Fig. 5 ML phylogenetic tree based on *COI* (a) and *18S rRNA* (b) gene sequences of zooplankton

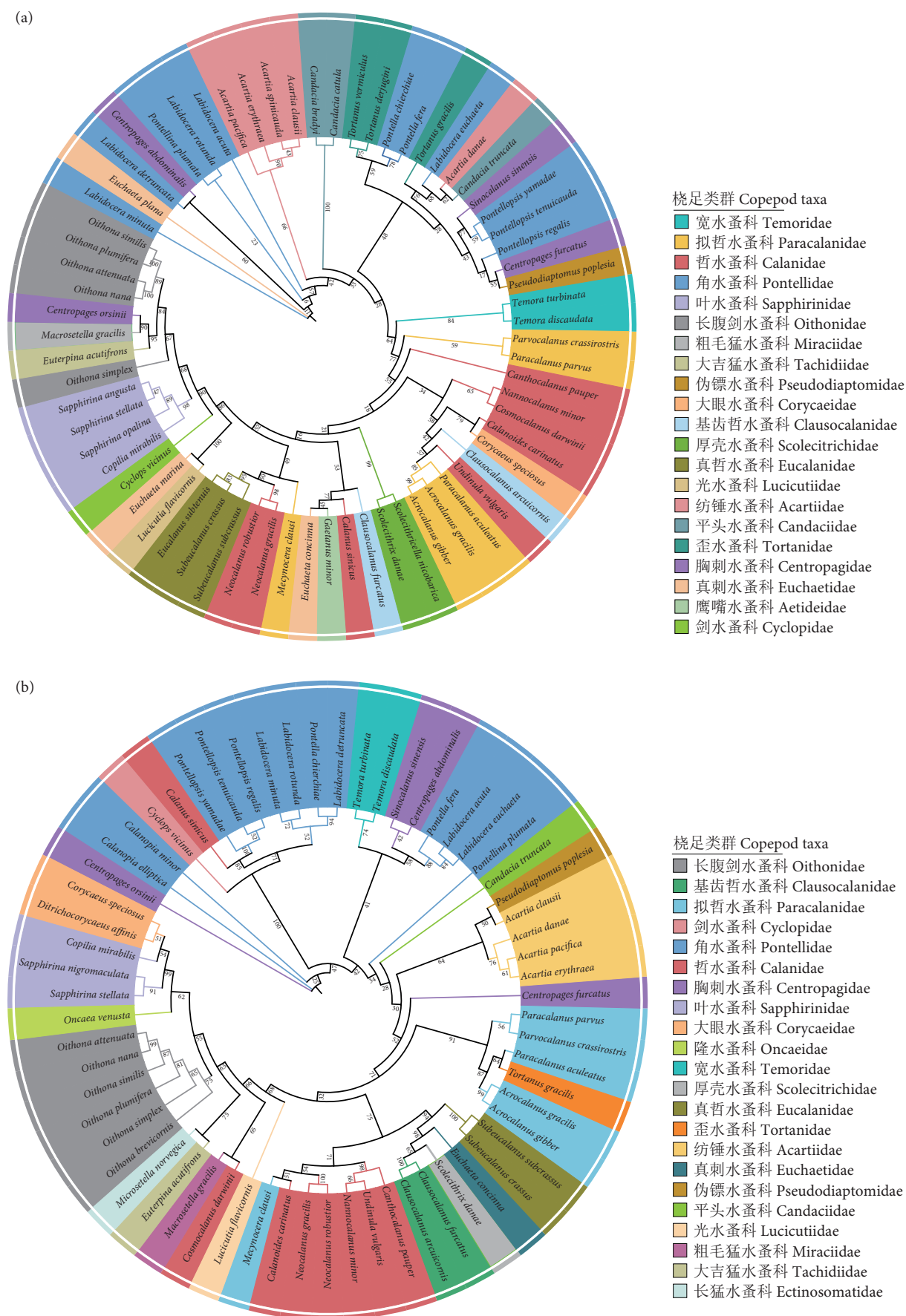


图6 基于 COI (a) 和 18S rRNA (b) 基因序列构建的桡足类 ML 系统发育树

Fig. 6 ML phylogenetic tree based on COI (a) and 18S rRNA (b) gene sequences of copepods

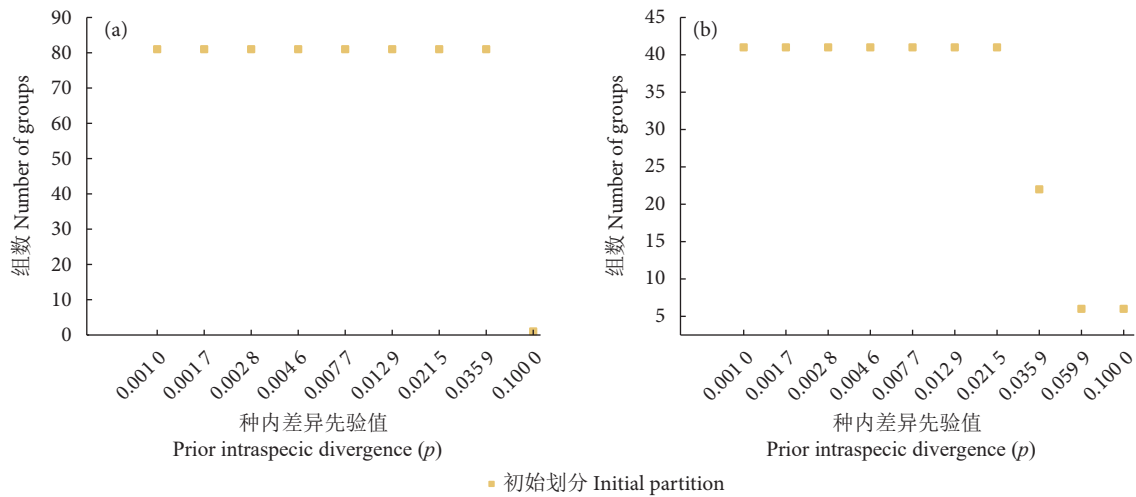


图7 基于浮游动物 *COI* (a) 和 *18S rRNA* (b) 基因序列的 ABGD 分析

Fig. 7 ABGD analysis based on *COI* (a) and *18S rRNA* (b) gene sequences of zooplankton

目、管水母目、花裸鳃目和筐水母目的物种则聚为一个分类单元。哲水蚤目和剑水蚤目之间的部分物种不能得到有效划分, 两个目的物种聚为一个分类单元, 与 ML 系统发育树的分类结果存在差异。

68 条桡足类 *COI* 基因序列经分析可被划分为 57 个分类单元 (图 8-a), 有 11 个物种聚为一个分类单元, 与对全部浮游动物的 *COI* 基因序列进行分类单元划分的结果相一致, 表明 *COI* 基因序列能有效区分桡足类与其他浮游动物类群。63 条桡足类 *18S rRNA* 基因序列可被划分为 53 个分类单元

(图 8-b), 其中, 隆哲水蚤属、基齿哲水蚤属、新哲水蚤属和长腹剑水蚤的物种分别聚为属的分类单元。结果表明, 基于 *18S rRNA* 基因序列, ABGD 方法对仅包含单一浮游动物类群的序列的物种分辨率要高于包含多个浮游动物类群的序列。这可能是因为不同浮游动物类群的 *18S rRNA* 基因序列存在较大差异, 在对序列进行对齐剪切时, 各浮游动物类群可变区被部分裁剪, 使得基于条形码间隔 (Barcode gap) 进行划分的 ABGD 方法不能对各物种进行有效划分。

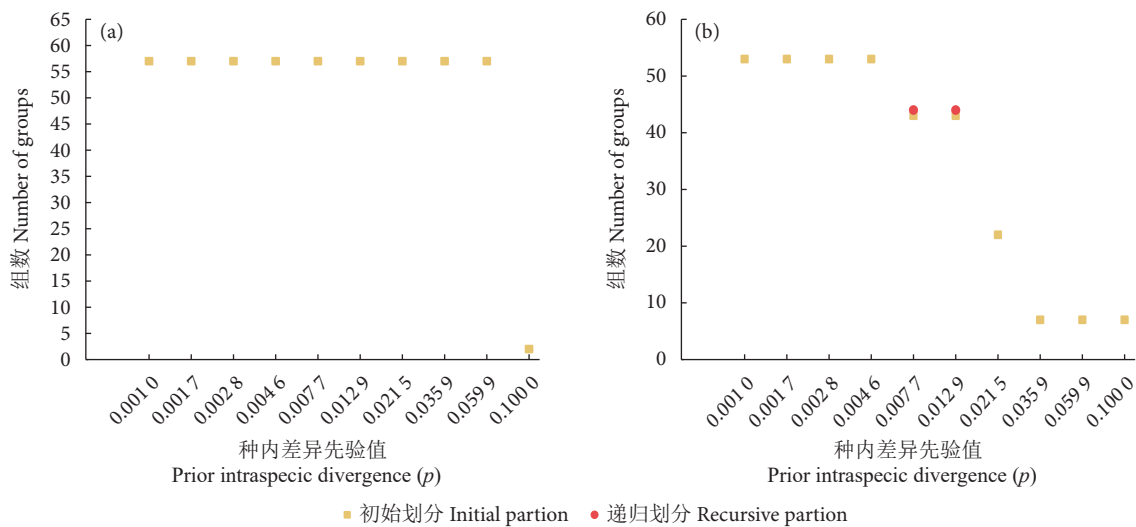


图8 基于桡足类 *COI* (a) 和 *18S rRNA* (b) 基因序列的 ABGD 分析

Fig. 8 ABGD analysis based on *COI* (a) and *18S rRNA* (b) gene sequences of copepods

3 讨论

3.1 条形码数据库的物种覆盖度

本研究共获得了 124 种浮游动物的 DNA 条形

码序列, 占舟山海域统计浮游动物的 80%。在 GenBank 数据库中, 有 31 种浮游动物没有相应的 DNA 条形码信息, 主要为桡足类 (10 种)、十足目

(4 种) 及糠虾目 (4 种) 等。其中, 常见的桡足类优势种如拟长腹剑水蚤 (*O. similis*) 和背针胸刺水蚤 (*C. dorsispinatus*) 的条形码信息也尚未包含, 这可能导致通过形态学鉴定和 DNA 分子鉴定 2 种方法计算的浮游动物优势类群出现分歧。在本研究中, 磷虾目的条形码覆盖率为 100%, 但糠虾目的条形码覆盖率仅为 33.33%。这可能是因为糠虾各物种的外部形态具有明显差异, 采用传统的形态学方法即可完成分类鉴定; 而磷虾由于各物种间外部形态相似^[36], 需要利用分子方法来实现准确的物种鉴定。另一方面, 一些浮游动物物种的 DNA 条形码标记基因的测序可能已经完成, 但出于隐私等原因尚未上传到公共数据库, 导致公共数据库中某些浮游动物序列信息的缺失。

为分析本研究构建的浮游动物条形码数据库在物种鉴定和多样性评估方面, 相较于其他公共数据库的优势和不足, 选取国际生命条形码数据系统 (Barcode of Life Data Systems, BOLD) (<https://v4.boldsystems.org/>) 和 MetaZooGene Barcode Atlas and Database (MZGdb) (<https://metazooengine.org/data-base/>) 这 2 个公共数据库, 对其在舟山海域浮游动物的物种覆盖度进行评估。因为海洋浮游动物普查 (Census of marine zooplankton, CMarz) 项目已于 2010 年结束, 相关浮游动物序列数据已不再更新, 故不再对其进行比较分析。基于编制的舟山海域浮游动物物种名录, 在 BOLD 和 MZGdb 数据库上检索相应的 COI 和 18S rRNA 基因序列, 结果表明, 在 BOLD 数据库上, 可检索到 119 种浮游动物的 COI 基因序列, 物种覆盖度达 76.77%, 高于本数据库 70.32% 的覆盖度; 但 BOLD 数据库中仅包含 28 种浮游动物的 18S rRNA 基因序列, 覆盖度仅为 18.06%, 低于本数据库的 58.71%。在 MZGdb 数据库中, 浮游动物 COI 和 18S rRNA 基因序列的物种覆盖度则分别为 73.55% 和 50.97%, 其在 18S rRNA 基因序列上的物种覆盖度较 BOLD 数据库更高。BOLD 和 MZGdb 这 2 个公共数据库在舟山海域浮游动物 COI 基因序列的物种覆盖度方面均高于本研究构建的数据库, 但也存在很多采自同一地区的同一物种的重复序列, 且在 18S rRNA 基因序列上的物种覆盖度均低于本数据库。BOLD 和 MZGdb 数据库中的浮游动物序列数据, 除了大部分来源于 NCBI GenBank 数据库外, 还有部分来自

未将序列数据上传到 GenBank 的科研院所 (如 Centre for Biodiversity Genomics)。因此, 未来可利用 BOLD 和 MZGdb 数据库中特有的浮游动物 COI 及 18S rRNA 基因序列来进一步扩充本数据库, 以提高浮游动物 DNA 条形码序列的物种覆盖度。

3.2 2 种 DNA 条形码基因在浮游动物鉴定中的适用性分析

COI 和 18S rRNA 基因已分别应用于浮游动物的物种鉴定^[37]。但目前尚未针对这 2 种 DNA 条形码基因在浮游动物的遗传多样性、物种界定的可行性上进行详细的报道。本研究从遗传多样性、系统发育关系和物种界定 3 个方面分析 COI 和 18S rRNA 基因作为 DNA 条形码基因在浮游动物种类鉴定中的适用性和局限性。首先, 基于 COI 和 18S rRNA 基因序列构建 ML 系统发育树, 结果发现在目和科等较高分类阶元上, 同一目或同一科的物种能以较高的支持率聚为单系支, 哲水蚤科、角水蚤科、基齿哲水蚤科等科的单系支的支持度均在 80% 以上, 表明这 2 种基因均能有效区分不同浮游动物类群。然而, 在 ML 系统发育树中, 也存在隶属不同目和科的物种聚为一支的情况, 如淡水水母目和花裸鳃目、哲水蚤目和端足目、胸刺水蚤科和伪镖水蚤科之间个别物种的聚类, 这些分歧支的支持度均低于 60%, 表明这些物种间可能存在较近的亲缘关系, 其分类地位仍需进一步探究。此外, 淡水水母目和花裸鳃水母目之间的分歧现象仅出现在基于 COI 基因序列构建的发育树中, 在基于 18S rRNA 基因序列构建的发育树中并未观察到。这可能是由于 COI 基因序列在不同浮游动物类群中进化速率存在差异, 导致远缘类群在系统发育树上异常聚类, 表明对浮游动物的系统发育关系分析不能仅依赖单一基因, 需要依赖多个基因及形态学特征共同分析。

在本研究中, 使用 K2P 模型计算浮游动物 COI 和 18S rRNA 基因序列的遗传距离, 而浮游动物 COI 基因序列的 A+T 含量在 60% 以上, 存在 A+T 偏倚性, 可能会对基于 K2P 模型计算的遗传距离结果产生影响; 18S rRNA 基因序列的 A+T 含量仅为 52%, 偏倚性较低。针对浮游动物 COI 基因序列中的 A+T 偏倚性, 可考虑使用 Tamura-Nei 模型计算种间及种内遗传距离, 该模型在计算时考虑了 2 类核苷酸之间的转换/颠换率不同的情况。由

于 K2P 模型已多次用于浮游动物如水母类、磷虾等 *COI* 基因序列遗传距离的计算^[36,38], 为了研究的一致性以及更好地比较 *COI* 和 *18S rRNA* 基因序列在浮游动物物种界定上的可行性, 本研究使用 K2P 模型进行种内及种间遗传距离的计算。在本研究中, 桡足类的 *COI* 和 *18S rRNA* 基因序列的种间平均遗传距离均为种内平均遗传距离的 10 倍以上, 满足作为 DNA 条形码基因的要求, 表明这 2 种基因均适用于桡足类的物种鉴定, 但 *18S rRNA* 基因序列在种水平上没有形成明显的条形码间隔 (图 4), 在桡足类近缘种上的分辨率要低于 *COI* 基因序列。究其原因, *18S rRNA* 基因在浮游动物中具有很高的遗传均质性, 其序列变异性要低于 *COI* 基因。首先, *18S rRNA* 基因编码的 *18S rRNA* 是核糖体小亚基重要的组成成分, 对 mRNA 识别及维持核糖体的结构稳定性至关重要。对 *18S rRNA* 结构稳定性的要求使得 *18S rRNA* 基因高度保守, 突变率低。而 *COI* 基因为蛋白质编码基因, 在该序列上的突变如为同义突变, 并不会影响编码蛋白质的正常功能, 使得 *COI* 基因对基因突变的耐受度要高于 *18S rRNA* 基因。其次, 浮游动物的 *COI* 基因比 *18S rRNA* 基因具有更高的进化速率^[39], 使其具有更高的种间遗传差异, 更适用于种水平的物种鉴定。

同时, ABGD 分析和 ML 系统发育树的结果表明, *COI* 和 *18S rRNA* 基因序列在目水平上对水母类浮游动物有较高的分辨率, 但不适用于种水平的物种界定。虽然 *COI* 和 *18S rRNA* 基因序列在水母类群中分辨率较低, 但研究表明 *16S rRNA* 基因序列在水母类群上具有很高的分辨率^[40-41], 尤其是在属水平上。除了使用 *16S rRNA* 基因序列, 还可使用核糖体内转录间隔区 (Internal transcribed spacer, *ITS*) 基因序列来提高 DNA 条形码基因在浮游动物分类界定上的分辨率。王仁诚^[42] 在对我国近海水母类和端足类的系统发育研究中, 利用 *ITS* 基因序列对涉及的浮游动物类群进行了准确鉴定, 证明了 *ITS* 基因序列在浮游动物类群尤其是水母类物种鉴定上的适用性。因此, 可针对不同的浮游动物样本, 有针对性地选择合适的 DNA 条形码基因用于物种界定。

Hebert 等^[35] 提出物种界定 2% 的阈值标准, 即种内遗传距离应不超过 0.02。在本研究中, 基

于 *18S rRNA* 基因序列, 有 4 个桡足类物种的种内遗传距离大于 0.02; 基于 *COI* 基因序列, 则有 2 个桡足类物种的种内遗传距离大于 0.02。由于浮游动物中存在大量的隐存种, 桡足类和毛颚类等类群的种内遗传距离差异较大^[43], 甚至与种间遗传距离发生重叠; 且 *COI* 和 *18S rRNA* 基因在不同的类群中进化速率不一致。因此, 2% 的物种界定阈值标准不一定适用于全部的浮游动物类群。针对桡足类和毛颚类等类群, 可设定 3% 为初始阈值, 并分析 *28S rRNA*、*16S rRNA* 等其他基因在浮游动物种内及种间的遗传差异, 以进一步提高物种分辨率。

本研究表明, *18S rRNA* 基因对桡足类等大中型浮游动物在种水平上的分辨率要低于 *COI* 基因, 但林渊源等^[44] 的研究表明 *18S rRNA* 基因具有更高的物种覆盖率。截至 2025 年 1 月, 以 “ciliates” 作为关键词在 GenBank 数据库上进行检索时, 检索到的 *18S rRNA* 基因序列 (10 753 条) 远多于 *COI* 基因序列 (1 322 条), 表明 *18S rRNA* 基因较 *COI* 基因更常用于原生动物的物种鉴定。因此, 为提高 DNA 条形码方法在浮游动物物种鉴定中的准确性及覆盖率, 可以选择多基因标记的 DNA 条形码浮游动物物种分类鉴定方法, 以避免低估研究区域的浮游动物群落的物种多样性。Zhao 等^[45] 采用 *COI* 和 *18S rRNA* 2 种标记基因对三门峡水库浮游动物群落结构进行研究, 结果表明纤毛虫在群落组成中占据绝对优势。

本研究存在一定局限性, 所编制的舟山海域浮游动物物种名录仅涵盖桡足类等大中型浮游动物, 未包含原生动物等微型浮游动物, 未能全面反映该海域浮游动物群落结构特征。同时, GenBank 数据库中, 舟山海域浮游动物的 *COI* 及 *18S rRNA* 基因的覆盖率均未超过 80%, 无法分析部分物种间的系统发育关系。后续研究应继续开展浮游动物采集工作, 通过个体分离与形态学鉴定相结合的方法, 对样本进行 DNA 提取、PCR 扩增及高通量测序, 获取物种 DNA 条形码基因序列, 逐步提升舟山海域浮游动物 DNA 宏条形码数据库的物种覆盖率。

综上所述, 本研究证实了 *COI* 和 *18S rRNA* 基因在舟山海域浮游动物物种鉴定中的可行性, 这 2 种 DNA 条形码基因在科及目阶元上均可有效区分不同浮游动物类群。本研究同时表明, *COI* 基因对浮游动物近缘种的分辨率要高于 *18S rRNA* 基

因。鉴于 18S rRNA、COI 和 16S rRNA 等条形码基因在物种鉴定上各有优劣, 建议后续研究采用多基因 DNA 条形码分析方法, 以更全面地分析浮游动物群落组成, 降低多样性评估偏差。

4 结论

1) COI 和 18S rRNA 基因的种间平均遗传距离均为种内平均遗传距离的 10 倍以上, 符合作为 DNA 条形码基因的标准; 2) 基于 ABGD 分类和 ML 系统发育树分析结果, COI 和 18S rRNA 基因在目及科等较高分类阶元上能有效界定不同的浮游动物类群, 但在种水平上对水母类物种的分辨率较低; 3) 18S rRNA 基因具有较高的遗传均质性, 在种水平上的分辨率要低于 COI 基因; 4) 自建数据库在舟山海域浮游动物 18S rRNA 基因的物种覆盖度要高于 BOLD 和 MZGdb 数据库, 但仍须借助 BOLD 等公共数据库进一步提高 COI 基因的物种覆盖度。

参考文献:

- [1] 胡吉连, 杜晓琴. 舟山海域悬浮体的特征及输运机制[J]. 海洋地质与第四纪地质, 2020, 40(6): 39-48.
- [2] 朱梦华, 钱卫国. 舟山渔场渔业资源衰退原因及修复对策[J]. 农村经济与科技, 2022, 33(9): 79-82.
- [3] LEVINE S N, BORCHARDT M A, BRANER M, et al. The impact of zooplankton grazing on phytoplankton species composition and biomass in Lake Champlain (USA-Canada)[J]. J Great Lakes Res, 1999, 25(1): 61-77.
- [4] 孙军, 宋煜尧, 施义锋, 等. 近十年中国海洋生物多样性研究进展[J]. 生物多样性, 2022, 30(10): 181-195.
- [5] 王腾, 刘永, 全秋梅, 等. 广东江门市主要淡水河流浮游动物群落结构特征[J]. 南方水产科学, 2021, 17(4): 9-17.
- [6] NDAH A B, MEUNIER C L, KIRSTEIN I V, et al. A systematic study of zooplankton-based indices of marine ecological change and water quality: application to the European marine strategy framework directive (MSFD)[J]. Ecol Indic, 2022, 135: 108587.
- [7] 冯芸芝, 孙栋, 邵倩文, 等. DNA 宏条形码技术在海洋浮游动物多样性和生态学研究中的应用[J]. 生态学报, 2022, 42(21): 8544-8554.
- [8] 高养春, 李海涛, 王孝程, 等. 利用宏 DNA 条形码研究浮游动物多样性: 以鸭绿江口为例[J]. 生态学报, 2020, 40(11): 3822-3832.
- [9] 杨蓓, 蔡倩, 江嵩鹤, 等. 基于 eDNA 和形态学的九江市城市湖泊浮游动物分析[J]. 江西水利科技, 2024, 50(4): 288-294.
- [10] 刘山林. DNA 条形码参考数据集构建和序列分析相关的新兴技术[J]. 生物多样性, 2019, 27(5): 526-533.
- [11] CHIARELLO M, MCCAULEY M, VILLEGGER S, et al. Ranking the biases: the choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold[J]. PLoS One, 2022, 17(2): e0264443.
- [12] CHEN Q Y, ZOBEL J, VERSPOOR K. Duplicates, redundancies and inconsistencies in the primary nucleotide databases: a descriptive study[J]. Database, 2017, 2017: baw163.
- [13] FICETOLA G F, PANSU J, BONIN A, et al. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data[J]. Mol Ecol Resour, 2015, 15(3): 543-556.
- [14] 宋晨, 孟周, 王晓波, 等. 2019—2020 年夏季舟山海域浮游动物优势种生态位及其生态分化[J]. 海洋学报, 2022, 44(10): 127-139.
- [15] 胡亚丽. 舟山近海浮游动物群落结构及多样性特征的季节变化研究[D]. 舟山: 浙江海洋大学, 2022: 27-30.
- [16] 陆延, 程爱勇, 林波, 等. 舟山电厂附近海域夏季浮游动物群落结构及时空变化[J]. 安徽农业科学, 2022, 50(4): 74-78.
- [17] 胡亚丽, 梁君, 许永久, 等. 2019 年舟山近海浮游动物群落结构变化与水团的关系[J]. 水产科学, 2022, 41(5): 788-798.
- [18] 边佳胤, 余健涛, 谢挺, 等. 春、夏季舟山海域浮游动物分布特征、群落结构及其与环境因子的关系[J]. 海洋湖沼通报, 2020(2): 122-130.
- [19] 孟庆超, 赵淑江. 舟山东极岛海域浮游动物群落结构特征[J]. 农村经济与科技, 2019, 30(11): 63-66.
- [20] 朱明星, 俞存根, 许永久, 等. 舟山近岸海域秋季桡足类种类组成和数量分布[J]. 水产科学, 2017, 36(4): 514-518.
- [21] 何舟洋, 许永久, 俞存根, 等. 舟山近岸海域秋季浮游动物多样性与环境因子的关系[J]. 浙江海洋学院学报(自然科学版), 2017, 36(2): 144-150.
- [22] 覃涛, 俞存根, 陈小庆, 等. 舟山渔场及邻近海域桡足类种类组成和数量分布[J]. 自然资源学报, 2011, 26(5): 780-789.
- [23] TZAFESTA E, SACCOMANNO B, ZANGARO F, et al. DNA barcode gap analysis for multiple marker genes for phytoplankton species biodiversity in Mediterranean aquatic ecosystems[J]. Biology, 2022, 11(9): 1277.
- [24] 郭梦月. 基于 DNA 条形码技术的中草药鉴定、分子系统学及污染真菌多样性研究[D]. 北京: 北京协和医学院, 2021: 47-50.
- [25] 闫亚利, 张楠, 郭华阳, 等. DNA 条形码在篮子鱼科鱼类种类鉴定和系统进化分析中的应用[J]. 南方水产科学, 2019, 15(1): 100-105.
- [26] 王开杰, 徐永江, 柳学周, 等. DNA 条形码在鲷属鱼类物种鉴定和系统进化分析中的适用性[J]. 中国水产科学, 2022, 29(2): 171-183.
- [27] 刘青青, 董志军. 基于线粒体 COI 基因分析钩手水母的群体遗传结构[J]. 生物多样性, 2018, 26(11): 1204-1211.
- [28] 孙明媛, 朱锐, 孙晓晴, 等. 基于 COI 和 16S rRNA 的库页岛马珂蛤遗传多样性和分子系统进化研究[J]. 中国水产科学, 2018, 25(4): 891-901.
- [29] 杜明敏, 刘镇盛, 王春生, 等. 中国近海浮游动物群落结构及季

- 节变化[J]. 生态学报, 2013, 33(17): 5407-5418.
- [30] 李超伦, 王敏晓, 程方平, 等. DNA 条形码及其在海洋浮游动物生态学研究中的应用[J]. 生物多样性, 2011, 19(6): 805-814.
- [31] TAKUMA N, KEIKO S. An extension of the Kimura Two-Parameter Model to the natural evolutionary process[J]. *J Mol Evol*, 2019, 87: 60-67.
- [32] COLLINS R A, BOYKIN L M, CRUICKSHANK R H, et al. Barcoding's next top model: an evaluation of nucleotide substitution models for specimen identification[J]. *Methods Ecol Evol*, 2012, 3(3): 457-465.
- [33] LETUNIC I, BORK P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. *Nucleic Acids Res*, 2021, 49(W1): W293-W296.
- [34] 刘若愚, 孙忠民, 姚建亭, 等. 中国近海重要生态建群红藻真江蓠的群体遗传多样性[J]. 生物多样性, 2016, 24(7): 781-790.
- [35] HEBERT P D N, CYWINSKA A, BALL S L, et al. Biological identifications through DNA barcodes.[J]. *Proc Biol Sci*, 2003, 270(1512): 313-321.
- [36] 湛微, 马春艳, 冯春雷, 等. 基于线粒体 *COI* 基因序列的磷虾目分子系统进化[J]. 中国水产科学, 2018, 25(4): 867-879.
- [37] FENG Y Z, SUN D, SHAO Q W, et al. *COI* metabarcoding better reveals the seasonal variations in the zooplankton community in the western Pacific warm pool[J]. *Ecol Indic*, 2023, 156(1): 111183.
- [38] 张瑛妮, 郑连明, 何劲儒, 等. 基于线粒体 *COI* 和 *16S* 片段序列的北部湾北部水螅水母 DNA 条形码分析[J]. 生物多样性, 2015, 23(1): 50-60.
- [39] WU S, JIE X, YU Y H. Taxonomic resolutions based on *18S rRNA* genes: a case study of subclass copepoda.[J]. *PLoS ONE*, 2017, 10(6): e0131498.
- [40] ZHENG L M, HE J R, LIN Y S, et al. *16S rRNA* is a better choice than *COI* for DNA barcoding hydrozoans in the coastal waters of China[J]. *Acta Oceanolog Sin*, 2014, 33(4): 55-76.
- [41] YE L J, PENG S J, MA Y Q, et al. Biodiversity and distribution patterns of blooming jellyfish in the Bohai Sea revealed by eDNA metabarcoding[J]. *BMC Ecol Evol*, 2024, 24(1): 37.
- [42] 王仁诚. 中国近海常见水母类和钵类的 DNA 条形码分析及系统发生研究[D]. 青岛: 中国科学院海洋研究所, 2010: 37-40.
- [43] 王仁诚, 王敏晓, 程方平, 等. 基于线粒体 *cox1* 片段序列胶州湾浮游动物物种组成分析[J]. 海洋与湖沼, 2011, 42(5): 711-721.
- [44] 林渊源, 赵峥. 环境 DNA 技术在水生入侵生物监测中的应用[J]. 生态毒理学报, 2021, 16(6): 1-12.
- [45] ZHAO L, ZHANG X, XU M Y, et al. DNA metabarcoding of zooplankton communities: species diversity and seasonal variation revealed by *18S rRNA* and *COI*[J]. *PeerJ*, 2021, 9: e11057.