

DOI: 10.12264/JFSC2020-6003

长臂虾科线粒体基因组结构与系统进化分析

朱陇强^{1, 2, 3}, 朱志煌^{2, 3}, 林琪^{2, 3}, 朱雷宇^{2, 3}, 王定全¹, 王健鑫¹

1. 浙江海洋大学海洋微生物分子生态与应用实验室, 浙江 舟山 316000;

2. 福建省水产研究所, 福建 厦门 361013;

3. 福建省海洋生物增殖与高值化利用重点实验室, 福建 厦门 361013

摘要: 综合分析了长臂虾科(Palaemonidae)14个物种线粒体基因组的全序列, 发现长臂虾科线粒体基因组的长度为15396~15967 bp, A+T 含量为 58.97%~69.09%; 长臂虾科内不同属线粒体基因组的基因排列各不相同, 与十足目(Decapoda)线粒体基因组的原始排列相比, 除沼虾属(*Macrobrachium*)的基因排列顺序未发生改变外, 长臂虾属(*Palaemon*)、叶颚虾属(*Hymenocera*)和贝隐虾属(*Anchistus*)的基因排列顺序均发生了不同程度的变化, 此外贝隐虾属(*Anchistus*)还出现基因缺失现象; 长臂虾属和沼虾属线粒体 13 个蛋白编码基因(protein-coding genes, PCGs)的 $K_a/K_s < 1$, 显示出很强的纯化选择; 长臂虾科 13 PCGs 在密码子的使用中, 被编码的氨基酸均体现相似的偏好性; 在线粒体基因组的差异位点分析中, 发现 *nd5*、*nd4* 和 *nd2* 基因为理想的分子标记。基于 13 PCGs 系统发育结果表明, 长臂虾属与沼虾属互为姊妹关系, 叶颚虾属和贝隐虾属单独聚为一大支。分子钟的估算结果推测长臂虾科物种可能起源于二叠纪, 随后进一步分化为具有现代表征的长臂虾种类。本研究为快速鉴定长臂虾科生物提供了可靠的分子标记, 并为分析长臂虾科物种遗传多样性提供理论基础。

关键词: 长臂虾科; 线粒体基因组; 结构特征; 系统发育

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2021)07-0852-11

长臂虾科(Palaemonidae)属甲壳动物亚门(Crustacea)、软甲纲(Malacostraca)、十足目(Decapoda)、真虾下目(Caridea); 包含众多经济价值很高的物种, 是目前真虾类中最大的科级分类单元之一^[1]。全世界共计长臂虾科物种约 981 种, 广泛分布于全球海洋和淡水环境中^[2]。自长臂虾科系统分类确立后, 其涵盖的范围一直颇受争议。长臂虾科生物在形态、生境和生活习性上差异很大, 仅靠形态特征或单基因研究其物种分类和系统进化是远远不够的^[1]。

近年来, 随着基因组学技术的不断突破, 长臂虾科内各物种间的系统发育研究也逐步从形态学层次进入基因组层次。虾类线粒体基因组

(mitogenome)结构与其他脊椎动物相似, 呈闭合环状, 通常由 37 个编码基因、1 个控制区(CR)和基因与基因间的非编码序列组成。虾类线粒体基因组作为核外遗传物质, 以结构简单、基因排列紧凑、基因组小、进化速度快, 具有母系遗传且易于扩增测序, 已被越来越多的科研人员作为研究虾类物种鉴别、分类及系统发育的理想材料^[3]。目前, 有关长臂虾科物种线粒体全基因组的研究发展迅速, 至今 GenBank 数据库中已公布 18 个长臂虾科物种的线粒体全基因组序列, 剔除 3 个基因组序列存在部分缺失的物种和 1 个基因组序列显示为待验证状态的物种, 剩余 14 个全基因组序列被用于开展分子系统学的相关研究。

收稿日期: 2020-10-23; 修订日期: 2020-12-08.

基金项目: 国家重点研发计划项目(2019YFD0901305); 福建省海洋与渔业结构调整专项(2020HYJG01, 2020HYJG08); 舟山市科技局项目(2019C21011).

作者简介: 朱陇强(1992-), 男, 硕士研究生, 研究方向为虾类分子系统学. E-mail: lqiangzhu@sina.com

通信作者: 王健鑫, 教授, 主要从事海洋微生物生态研究. E-mail: jxwang@zjou.edu.cn

在本研究中,通过对 14 个长臂虾科物种线粒体全基因组序列的结构特征进行比较分析,并利用 13 PCGs 构建系统发育树探讨各物种之间的系统发育关系,为其寻找合适的分子标记、评估种群遗传多样性提供参考。

1 材料与方法

1.1 数据的获取

从 NCBI (<http://www.ncbi.nlm.nih.gov/Genbank/>)

公共数据库中检索到 14 种长臂虾科线粒体基因组全序列,其中未检索到中文名称的物种直接用拉丁名表示,物种相关的信息见表 1。

1.2 选择压力分析

为了更好地了解长臂虾属和沼虾属物种在 DNA 水平的进化和选择作用,采用 K_aK_s calculator 2.0^[4]软件计算它们的同义替换率(K_s)和非同义替换率(K_a),计算方法采用最大似然法中的模型平均法,遗传密码子选择无脊椎动物线粒体密码子。

表 1 长臂虾科线粒体基因组的基本特征
Tab. 1 Basic characteristics of mitogenome of Palaemonidae

分类地位 classification status	种名 species	GenBank 接收号 GenBank accession no.	长度/bp length	AT 含量/% AT content
沼虾属 <i>Macrobrachium</i>	<i>Macrobrachium bullatum</i>	NC_027602	15 774	64.58
	<i>Macrobrachium lanchesteri</i>	NC_012217	15 694	67.09
	日本沼虾 <i>Macrobrachium nipponense</i>	NC_015073	15 806	66.04
	罗氏沼虾 <i>Macrobrachium rosenbergii</i>	NC_006880	15 772	62.27
长臂虾属 <i>Palaemon</i>	安氏白虾 <i>Palaemon annandalei</i>	NC_038117	15 716	64.08
	<i>Palaemon capensis</i>	NC_039373	15 925	69.09
	葛氏长臂虾 <i>Palaemon gravieri</i>	NC_029240	15 740	67.44
	黄点玻璃虾 <i>Palaemon serenus</i>	NC_027601	15 967	58.97
	中华小长臂虾 <i>Palaemon sinensis</i>	NC_045090	15 955	66.61
	脊尾白虾 <i>Palaemon carinicauda</i>	NC_012566	15 730	63.60
	<i>Palaemon serratus</i>	NC_050266	15 758	66.77
	<i>Palaemon adspersus</i>	NC_050168	15 736	66.83
	贝隐虾属 <i>Anchistus</i>	NC_046034	15 396	68.40
叶颚虾属 <i>Hymenocera</i>	小丑虾 <i>Hymenocera picta</i>	NC_039631	15 786	65.68

1.3 基因变异位点分析

基因变异位点分析包含 2 个数据群,即长臂虾属类线粒体基因组和沼虾属类线粒体基因组。通过 MEGA 5.0 软件^[5]对 2 个数据群线粒体基因组编码基因和核糖体 RNA 基因进行多重序列比对,再用 DnaSP 6^[6]软件分析 2 个数据群线粒体基因组的基因变异位点。

1.4 遗传距离分析

将长臂虾科 14 个物种线粒体基因组的 13 PCGs 通过 MEGA 5.0 软件^[5]进行多重序列比对,用 Kimura 双参数法计算遗传距离,再利用 TBtools 软件^[7]绘制热谱图(heatmap)。

1.5 密码子使用模式

去除长臂虾科 14 个物种线粒体基因组 13

PCGs 的终止密码子,将核苷酸序列串联,利用 MEGA 5.0 软件^[5]计算密码子的选择数。

1.6 系统发育分析

使用 MEGA 5.0 软件对长臂虾科 14 个物种线粒体基因组的 13 PCGs 进行整理比对,并选取进化较为原始的匙指虾属物种 *Halocaridina rubra* 线粒体基因组的 13 PCGs 一起用于系统发育分析。比对后的序列通过 Gblocks^[8]去除不可靠的对齐序列。用 jModelTest 软件^[9]估算最适的核苷酸替代模型,根据赤池信息准则(AIC),GTR+I+G (-lnL=97214.4297)为最佳的替代模型。系统发育分析使用最大似然法(maximum likelihood: ML)和贝叶斯法(Bayesian inference: BI)来构建系统发育树。其中 ML 使用 PhyML 3.1^[10]构建,自展值(Bootstrap=

1000)评估节点可靠性,以 BP 表示,其他参数为默认值。BI 分析使用 MrBayes v3.2.6^[11]构建,运行世代数为 1000000,每 100 时代在屏幕上显示一次,每 100 个世代将取样的树保存,同时运行 4 个马尔可夫链,包括 3 条热链和 1 条冷链,测算贝叶斯后验概率,以 PP 表示。

1.7 分歧时间估计

为了避免影响估算长臂虾科各物种分化时间的准确性,去除系统发育分析中出现长枝现象的 *Anchistus australis* 的 13 PCGs,其余物种的 13 PCGs 串联序列均参与分歧时间估算。所用软件为 BEAST v. 2.6.3^[12],采用 GTR+I+G 模型和 Yule 模型先验信息进行分析,分子钟选择宽松分子钟模型。校准节点以沼虾属的分化时间为标准^[13]。进行马尔可夫分析的运行链长为 10000000,每运行 1000 次抽样一次,最后弃掉 50%的老化链。

2 结果与分析

2.1 长臂虾科线粒体基因组的特征及基因排列

由表 1 可知,长臂虾科 14 个物种线粒体基因组的长度在 15396 bp(*Anchistus australis*)和 15967 bp(黄点玻璃虾)之间。14 个物种线粒体基因组 A+T 含量最低的为 58.97%(黄点玻璃虾),最高的为 69.09% (*Palaemon capensis*)。

在基因重排方面,沼虾属内 4 个物种(*Macrobrachium bullatum*、*Macrobrachium lanchesteri*、日本沼虾、罗氏沼虾)线粒体基因组的基因排列顺序完全相同;长臂虾属内 8 个物种(安氏白虾、*Palaemon capensis*、葛氏长臂虾、黄点玻璃虾、中华小长臂虾、脊尾白虾、*Palaemon serratus*、*Palaemon adspersus*)线粒体基因组的基因排列顺序也完全相同(图 1)。与十足目线粒体基因组的原始排列相比,沼虾属 4 个物种的线粒体基因排列顺序未发

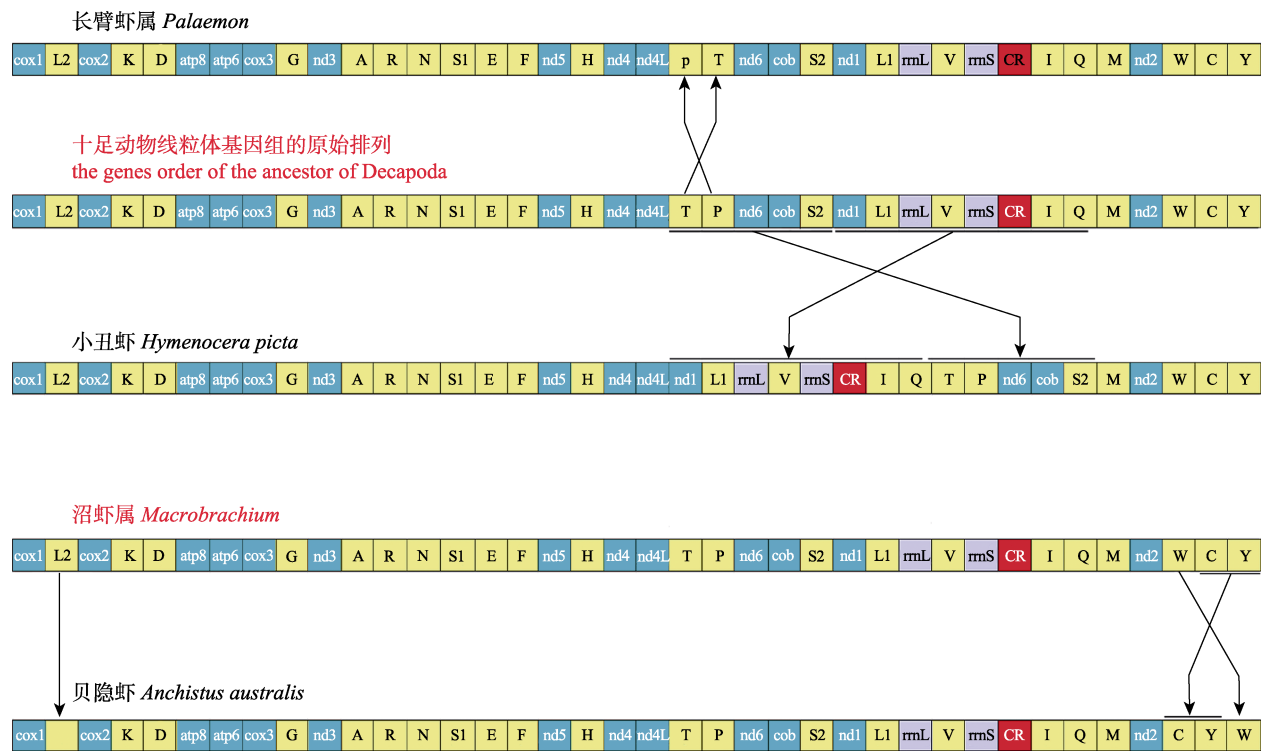


图 1 长臂虾科线粒体基因组图谱

长臂虾属: 安氏白虾、*Palaemon capensis*、葛氏长臂虾, 黄点玻璃虾、中华小长臂虾、脊尾白虾、*Palaemon serratus* 和 *Palaemon adspersus*; 沼虾属: *Macrobrachium bullatum*、*Macrobrachium lanchesteri*、日本沼虾和罗氏沼虾。
红色字体表示线粒体基因排列顺序一致。

Fig. 1 Mitogenome map of Palaemonidae

The *Palaemon* include *Palaemon annandalei*, *Palaemon capensis*, *Palaemon graviera*, *Palaemon serenus*, *Palaemon sinensis*, *Palaemon carinicauda*, *Palaemon serratus*, and *Palaemon adspersus*. The *Macrobrachium* include *Macrobrachium bullatum*, *Macrobrachium lanchesteri*, *Macrobrachium nipponense*, and *Macrobrachium rosenbergii*.
Red font indicates the same order of mitochondrial genes.

生变化, 长臂虾属物种的两个 tRNA 基因(*trnT* 和 *trnP*)位置均发生了易位, 小丑虾线粒体基因组的两个基因区块(*trnT-trnP-nad6-cob-S2* 和 *nad1-trnL1-rrnL-trnV-rrnS-CR-trnI-trnQ*)发生了易位, *Anchistus australis* 线粒体基因组的 *trnW* 基因与基因区块(*trnC-trnY*)发生易位且 *trnL2* 基因缺失(图 1)。

2.2 长臂虾科物种线粒体基因组的选择压力

长臂虾属内物种 13 PCGs 的 K_a/K_s 在 0.009(*cox1*) 和 0.122(*atp8*)之间, 沼虾属内物种 13 个编码基因的 K_a/K_s 在 0.007(*cox1*)和 0.050(*atp8*)之间, 均小于 0.2, 表现出很强的纯化选择(图 2)。在长臂虾属中, K_a/K_s 值最低的为 *cox1* 基因(平均为 0.009), 其次为 *cox2*、*cox3*、*cob* 和 *nd1* 基因(平均值为 0.016、0.019、0.019 和 0.023), 说明在长臂虾属中, 这些基因承受较强的选择压力; 而 K_a/K_s 值最高的为 *atp8* 基因(平均为 0.122), 其次是 *nd6*、*nd2*、*nd3* 和 *nd4L* 基因(分别为 0.097、0.076、0.051 和 0.050), 说明在长臂虾属中这些基因承受较弱的选择压力(图 2a)。13 个基因中 K_a/K_s 标准差最高的是 *nd6*

基因(0.016), 其次是 *atp8*、*nd3* 和 *nd2* 基因(分别为 0.015、0.013 和 0.011), 说明长臂虾属中这些基因存在较强的种间选择压力差异; K_a/K_s 标准差最低的是 *cox1* 基因(0.001), 其次是 *cob*、*cox2*、*cox3* 和 *nd1* 基因(分别为 0.002、0.003、0.003 和 0.003), 说明这些基因的种间选择压力差异较小。

在沼虾属中, K_a/K_s 值最低的为 *cox1* 基因(平均为 0.007), 其次是 *cob*、*nd1*、*nd4L* 和 *nd5* 基因(平均值为 0.008、0.009、0.009 和 0.012), 说明在沼虾属中, 这些基因承受较强的选择压力; 而 K_a/K_s 值最高的为 *atp8* 基因(平均为 0.050), 其次是 *nd6*、*nd3*、*nd2* 和 *atp6* 基因(平均值为 0.044、0.027、0.023 和 0.022), 说明在沼虾属中这些基因承受较弱的选择压力(图 2b)。13 个基因中 K_a/K_s 标准差最高的是 *atp8* 基因(0.011), 其次是 *nd6* 和 *nd3* 基因(分别为 0.010 和 0.008), 说明沼虾属中这些基因存在较强的种间选择压力差异; K_a/K_s 标准差较低的是 *cox1*、*nd4*、*nd5* 和 *cob* 基因(分别为 0.001、0.001、0.001 和 0.001), 说明这些基因的种间选择压力差异较小。

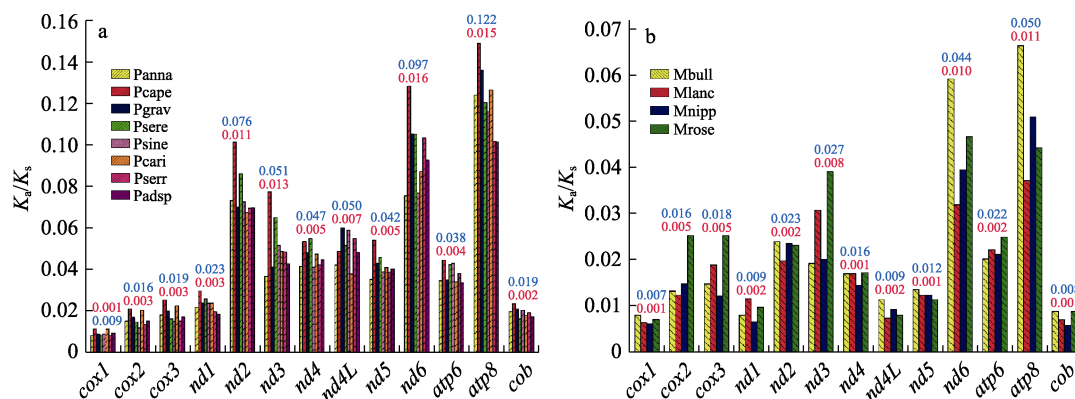


图 2 长臂虾属和沼虾属物种线粒体基因选择压力分析

a: 长臂虾属, 种名缩写为: 安氏白虾(Panna), *Palaemon capensis* (Pcape), 葛氏长臂虾(Pgrav), 黄点玻璃虾(Psere), 中华小长臂虾(Psine), 脊尾白虾(Pcari), *Palaemon serratus* (Pserr), *Palaemon adspersus* (Padsp); b: 沼虾属, 种名缩写为: *Macrobrachium bullatum* (Mbull), *Macrobrachium lanchesteri* (Mlanc), 日本沼虾(Mnipp), 罗氏沼虾(Mrose).

Fig. 2 Selection pressure analysis of mitogenomes in the genus *Palaemon* and *Macrobrachium*

a: *Palaemon*, species abbreviation: *Palaemon annandalei* (Panna), *Palaemon capensis* (Pcape), *Palaemon graviera* (Pgrav), *Palaemon serenensis* (Psere), *Palaemon sinensis* (Psine), *Palaemon carinicauda* (Pcari), *Palaemon serratus* (Pserr), *Palaemon adspersus* (Padsp); b: *Macrobrachium*, species abbreviation: *Macrobrachium bullatum* (Mbull), *Macrobrachium lanchesteri* (Mlanc), *Macrobrachium nipponense* (Mnipp), *Macrobrachium rosenbergii* (Mrose).

2.3 长臂虾科物种的分子标记

由表 2 可知, 长臂虾属线粒体基因组的 15 个基因中, 8 个基因(*atp6*、*atp8*、*nd2*、*nd3*、*nd4*、*nd4L*、*nd5* 和 *nd6*)变异位点比例超过 50%, 6 个基

因(*cob*、*cox2*、*cox3*、*nd1*、*rrnL* 和 *rrnS*)变异位点比例为 40%~50%; *cox1* 基因最为保守, 变异位点的比例为 38.63%(表 2)。变异位点数最多的是 *nd5* 基因(943 个), 其次是 *nd4* 和 *nd2* 基因(分别为 723

个和 626 个), 这 3 个基因不仅变异位点比例高、基因长度长, 而且变异位点多, 因此可以作为长臂虾属理想的分子标记。

沼虾属线粒体基因组的 15 个基因中, 6 个基因(*atp6*、*atp8*、*nd2*、*nd3*、*nd5* 和 *nd6*)变异位点比例超过 30%, 其余 9 个基因(*cob*、*cox1*、*cox2*、*cox3*、*nd1*、*nd4*、*nd4L*、*rrnL* 和 *rrnS*)变异位点比例均低于 30%(表 3)。变异位点数最多的是 *nd5* 基

因(544 个), 其次为 *cox1*、*nd4* 和 *nd2* 基因(分别为 397 个、390 个和 383 个)。由于 *cox1* 基因变异位点比例较低, 单独利用该基因不足以分析各物种与群落之间的生物多样性。因此, 选择 *nd5*、*nd4* 和 *nd2* 基因作为沼虾属较为理想的分子标记。综上所述, 两属具有相同的分子标记基因, *nd5*、*nd4* 和 *nd2* 基因可作为长臂虾科理想的分子标记来分析长臂虾不同群体之间的遗传多样性。

表 2 长臂虾属线粒体基因组的基因变异位点分析
Tab. 2 Analysis of gene variation sites in the mitogenome of *Palaemon*

参数 parameter	基因名称 gene name														
	<i>atp6</i>	<i>atp8</i>	<i>cob</i>	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>	<i>nd1</i>	<i>nd2</i>	<i>nd3</i>	<i>nd4</i>	<i>nd4L</i>	<i>nd5</i>	<i>nd6</i>	<i>rrnL</i>	<i>rrnS</i>
总位点数* total number of sites	675	159	1134	1530	688	789	936	987	354	1333	300	1694	512	1269	771
保守位点数 number of invariable sites	318	57	614	939	389	443	480	361	158	610	142	751	158	694	416
变异位点数 number of variable sites	357	102	520	591	299	346	456	626	196	723	158	943	354	575	355
变异位点比例/% ratio of variable sites	52.89	64.15	45.86	38.63	43.46	43.85	48.72	63.42	55.37	54.24	52.67	55.67	69.14	45.31	46.04

注: *表示总位点数不包含缺失位点。
Note: * means the total number of sites does not include missing sites.

表 3 沼虾属线粒体基因组的基因变异位点分析
Tab. 3 Analysis of gene variation sites in the mitogenome of *Macrobrachium*

参数 parameter	基因名称 gene name														
	<i>atp6</i>	<i>atp8</i>	<i>cob</i>	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>	<i>nd1</i>	<i>nd2</i>	<i>nd3</i>	<i>nd4</i>	<i>nd4L</i>	<i>nd5</i>	<i>nd6</i>	<i>rrnL</i>	<i>rrnS</i>
总位点数* total number of sites	673	159	1132	1535	684	789	939	994	354	1335	300	1706	515	1273	790
保守位点数 number of invariable sites	456	101	816	1138	495	562	682	611	230	945	221	1162	317	935	621
变异位点数 number of variable sites	217	58	316	397	189	227	257	383	124	390	79	544	198	338	169
变异位点比例/% ratio of variable sites	32.24	36.48	27.92	25.86	27.63	28.77	27.37	38.53	35.03	29.21	26.33	31.89	38.45	26.55	21.39

注: *表示总位点数不包含缺失位点。
Note: * means the total number of sites does not include missing sites.

2.4 长臂虾科物种的遗传距离

在长臂虾属内, *Palaemon serratus* 与 *Palaemon adspersus* 的遗传距离最小(0.1767), *Palaemon capensis* 与黄点玻璃虾的遗传距离最大(0.4005), 8 个物种之间的遗传距离均值为 0.3340; 在沼虾属内, *Macrobrachium lanchesteri* 与罗氏沼虾的遗传距离最小(0.1861), *Macrobrachium bul-latum* 与 *Macrobrachium lanchesteri* 的遗传距离最大(0.2244), 4 个物种之间的遗传距离均值为

0.2108; 在长臂虾属与沼虾属之间, 日本沼虾与 *Palaemon capensis* 的遗传距离最小(0.3190), *Macrobrachium lanchesteri* 与黄点玻璃虾的遗传距离最大(0.3960), 两属间的遗传距离均值为 0.3486(图 3), 属间的遗传距离大于属内遗传距离, 这完全支持传统的属级分类; 在整个长臂虾科内, *Palaemon serratus* 与 *Palaemon adspersus* 的遗传距离最小(0.1767), 黄点玻璃虾与 *Anchistus australis* 的遗传距离最大(0.4861)。同时, 基于 13 个

蛋白质编码基因所构建的系统发育结果也完全支持遗传距离分析结果(图 3)。

2.5 长臂虾科物种 13 PCGs 密码子的使用

在长臂虾属中, 编码异亮氨酸(Ile)、亮氨酸(Leu1, Leu2)、苯丙氨酸(Phe)、缬氨酸(Val)、甘氨酸(Gly)和丙氨酸(Ala)的密码子数最多(>200), 编码精氨酸(Arg)、谷氨酸(Glu)、谷氨酰胺(Gln)、组氨酸(His)、赖氨酸(Lys)、半胱氨酸(Cys)和天冬氨酸(Asp)的密码子最少(<100), 其他各属(沼虾属、贝隐虾属和叶颚虾属)也遵循同样的规律(图 4),

酸(Gly)和丙氨酸(Ala)的密码子数最多(>200), 编码精氨酸(Arg)、谷氨酸(Glu)、谷氨酰胺(Gln)、组氨酸(His)、赖氨酸(Lys)、半胱氨酸(Cys)和天冬氨酸(Asp)的密码子最少(<100), 其他各属(沼虾属、贝隐虾属和叶颚虾属)也遵循同样的规律(图 4),

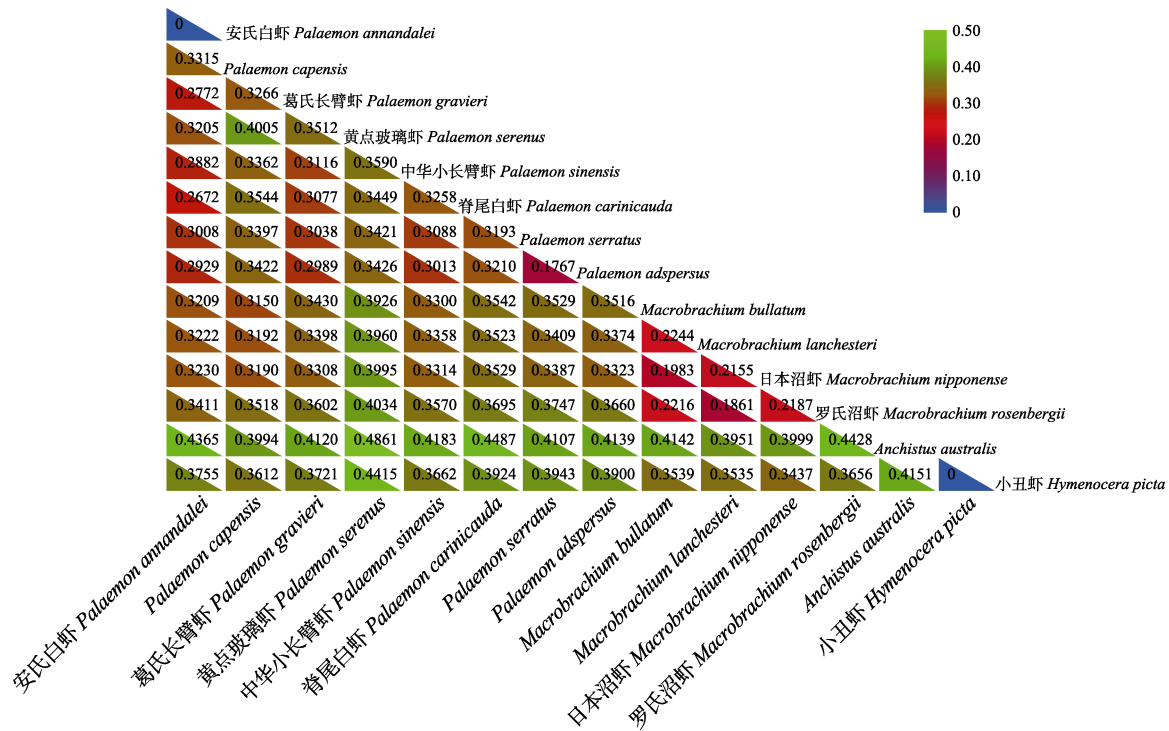


图 3 长臂虾科物种线粒体基因组 13 个蛋白编码基因的遗传距离
Fig. 3 The genetic distance of 13 protein-coding genes in the mitogenome of Palaemonidae species

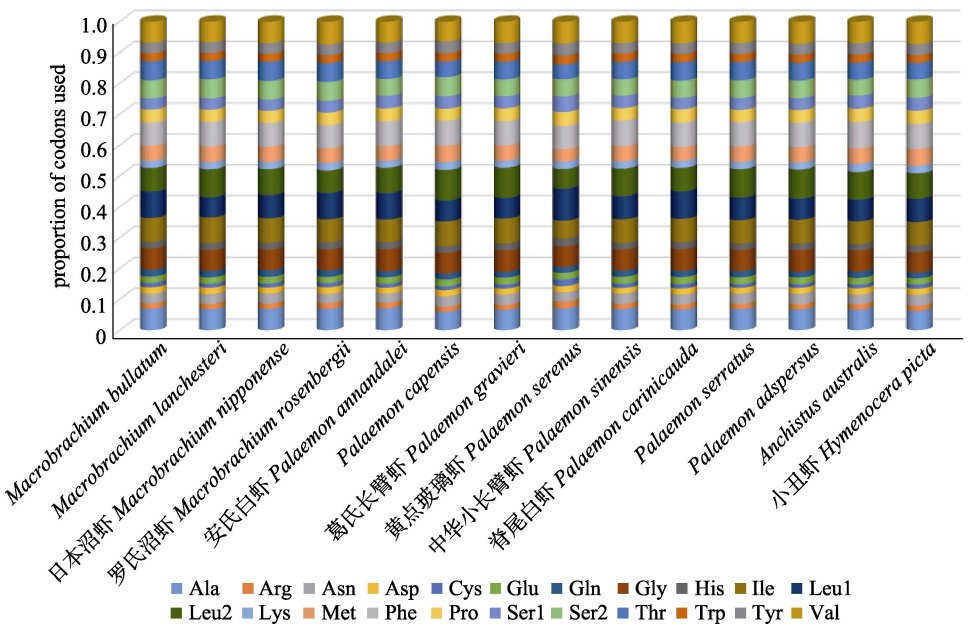


图 4 长臂虾科线粒体基因组密码子的使用模式
Fig. 4 Codon usage pattern in mitochondrial genome of Palaemonidae species

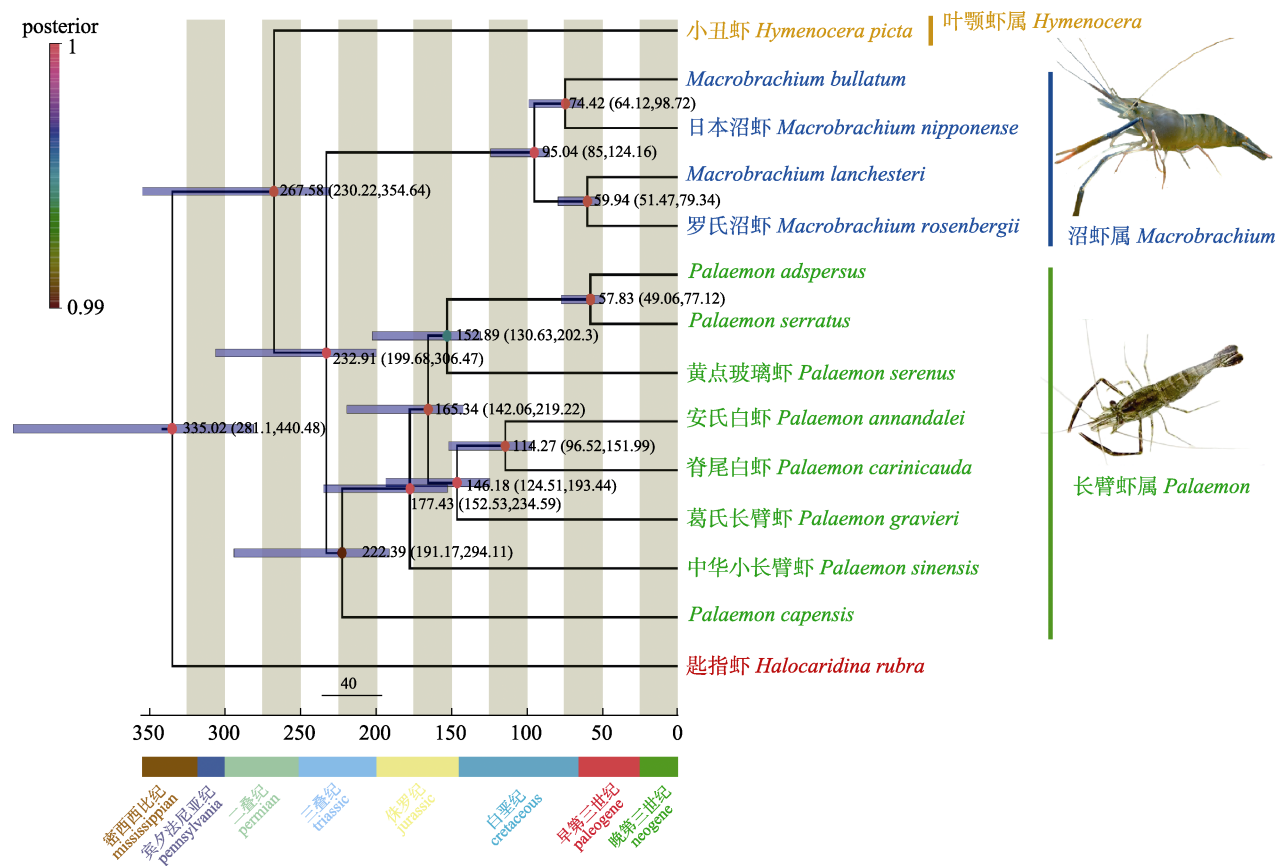


图 6 长臂虾科物种分歧时间估算
数字代表节点的中间值及 95% 的置信区间。
Fig. 6 Estimation of the divergence time of the species of Palaemonidae
The number represent posterior medians and 95% posterior interval value.

库现有的 14 个长臂虾科物种线粒体基因组数据进行分析发现,长臂虾科 14 个物种线粒体基因组的长度为 15396 (*Anchistus australis*)~15967 bp (黄点玻璃虾), A+T 含量为 58.97%~69.09%,符合正常十足目物种线粒体基因组的长度,AT 含量高也是十足目物种普遍存在的现象。在基因重排方面,基因重排在软甲纲动物线粒体基因组中是很常见的^[15]。与十足动物线粒体基因组的原始排列相比,沼虾属线粒体基因组排列顺序均未发生改变,说明沼虾属物种在进化中相当保守;8 个长臂虾属物种线粒体基因组的 *trnT* 基因和 *trnP* 基因位置均发生了易位, *trnT* 基因和 *trnP* 基因易位可能是长臂虾属共有的现象。作为海洋观赏性物种的小丑虾,与十足动物线粒体基因组的原始排列相比,两个基因区块发生了易位,这种基因区块的易位在虾类中很少见,但在蟹类中很常见^[16]。

Anchistus australis 线粒体基因组出现罕见的 *trnL2* 基因缺失现象,且 *trnW* 基因与基因区块 (*trnC-trnY*) 发生易位,目前已知柯氏绒铠虾 (*Shinkaia crosnieri*) 的线粒体基因组存在基因缺失现象^[17],太平洋磷虾 (*Euphausia pacifica*) 和南极磷虾 (*Euphausia superba*) 线粒体基因组的 *trnW* 基因与基因区块 (*trnC-trnY*) 均发生了易位^[18],对比长臂虾科中各属物种线粒体基因组基因的排列顺序,说明长臂虾科在基因排列中并不保守。同时,基因组重排模式也从基因组层次上支持 14 个长臂虾科物种系统发育关系(图 5)。

3.2 选择压力与分子标记

通常,在遗传学中,用 K_a/K_s 来判断是否有选择压力作用于这个蛋白编码基因。 $K_a/K_s > 1$ 、 $K_a/K_s = 1$ 和 $K_a/K_s < 1$ 分别表示基因受正选择 (positive selection)、中性进化 (neutral evolution) 和

纯化选择(purify selection)^[19]。选择压力分析显示,长臂虾属和沼虾属 13 PCGs 的 K_a/K_s 均小于 1,说明 13 个基因均受负选择的影响。两属 *cox1* 基因的 K_a/K_s 值最低,这在其他后生动物线粒体基因的研究中, *cox1* 基因的 K_a/K_s 值一般都为最小^[16,20-22]。另外值得注意的是,长臂虾属和沼虾属 *atp8* 基因的 K_a/K_s 值最高,说明该基因在 13 PCGs 中承受最小的选择压力,这可能是许多后生动物线粒体基因组 *atp8* 基因存在缺失的原因^[23-26]。此外,本文通过分析长臂虾属和沼虾属线粒体基因组的基因变异位点发现, *nd5*、*nd4* 和 *nd2* 基因作为长臂虾科理想的分子标记,这对长臂虾科物种遗传分析、分歧时间估算和地理分布研究至关重要^[20]。同时,这一结果也与申欣等^[15]的分析完全一致。

3.3 长臂虾科的系统发育与分歧时间

长臂虾科生物在形态、生境和生活习性上差异很大,有些物种可通过形态进行区分,而有些形态极其相似,仅靠形态特征或单基因研究是远远不够的,因此有必要利用分子手段对这些物种进行详细的系统发育分析。本文利用公共数据库 14 个长臂虾科物种线粒体基因组及 1 个匙指虾线粒体基因组做外群,通过最大似然法和贝叶斯法进行系统发育建树分析,结果显示这两种方法构建的系统发育树拓扑结构完全一致,长臂虾属和沼虾属分别作为一个独立单元且互为姊妹关系,这与前人基于线粒体基因组 13 个蛋白质编码基因的研究结果一致^[27]。小丑虾与 *Anchistus australis* 单独自聚为一大支,但 Sung 等^[28]基于 13 PCGs 构建的系统发育结果显示,叶颚虾属的小丑虾与沼虾属聚为一支且与长臂虾属互为姊妹关系。关于叶颚虾属的系统发育地位,Bracken 等^[29]基于 16S 和 18S 构建的系统发育树显示,叶颚虾属与 *Gnathophyllum* 聚为一支,与沼虾属不在同一个分支上,随后 De Grave 等^[30]基于 16S, H3 和 18S 构建的系统发育树也显示同样的结果,再次说明本文长臂虾科各物种系统发育关系的准确性。

根据对长臂虾科物种分歧时间的估算显示,长臂虾科物种可能是在 267.58 Mya 左右分化形成的,随后在不同的时间内形成适应不同生态环境的近缘物种。长臂虾属与沼虾属的分化时间发生

在 232.91 Mya (95%的 HPD 为 199.68~306.47 Mya),这与前人估算的 177.1 Mya 基本相近^[13];罗氏沼虾与沼虾 *Macrobrachium lanchesteri* 的分化时间发生在 59.94 Mya (95%的 HPD 为 51.47~79.34 Mya),这与前人估算的 46 Mya 基本相近^[13]。另外,长臂虾科与外群匙指虾科的分化时间发生在 335.02 Mya (95%的 HPD 为 281.1~440.48 Mya),这与生命进化时间信息公共知识库(<http://www.timetree.org/>)记录的 345 Mya 很接近。虽然本研究加入了许多物种新的分子分歧时间估算值,但总体分子钟估算结果与前人的研究结果基本相近,而且分歧时间的置信区间与前人的分析结果都有重叠,因此本文分子钟估算结果具有可信性。

参考文献:

- [1] Gan Z B, Li X Z. A preliminary phylogenetic analysis of the family Palaemonidae (Caridea: Palaemonoidea) based on mitochondrial 16S rRNA gene[J]. Marine Sciences, 2014, 38(7): 7-13. [甘志彬, 李新正. 基于 16S rRNA 基因片段的长臂虾科系统发育初步研究[J]. 海洋科学, 2014, 38(7): 7-13.]
- [2] De Grave S, Fransen C H J M. Carideorum catalogus: the recent species of the Dendrobranchiate, Stenopodidean, Procarididean and Caridean shrimps (Crustacea: Decapoda)[J]. Zoologische Mededelingen Leiden, 2011, 85(9): 195-589.
- [3] Xiao W H, Zhang Y P. Genetics and evolution of mitochondrial DNA in fish[J]. Acta Hydrobiologica Sinica, 2000, 24(4): 384-391. [肖武汉, 张亚平. 鱼类线粒体 DNA 的遗传与进化[J]. 水生生物学报, 2000, 24(4): 384-391.]
- [4] Zhang Z, Li J, Zhao X Q, et al. K_a/K_s Calculator: Calculating K_a and K_s through model selection and model averaging[J]. Genomics, Proteomics & Bioinformatics, 2006, 4(4): 259-263.
- [5] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. Molecular Biology and Evolution, 2011, 28(10): 2731-2739.
- [6] Rozas J, Ferrer-Mata A, Sánchez-Delbarrio J C, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets[J]. Molecular Biology and Evolution, 2017, 34(12): 3299-3302.
- [7] Chen C J, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Molecular Plant, 2020, 13(8): 1194-1202.
- [8] Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis[J]. Molecular Biology and Evolution, 2000, 17(4): 540-552.

- [9] Darriba D, Taboada G L, Doallo R, et al. jModelTest 2: More models, new heuristics and parallel computing[J]. *Nature Methods*, 2012, 9(8): 772.
- [10] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood[J]. *Systematic Biology*, 2003, 52(5): 696-704.
- [11] Ronquist F, Teslenko M, van der Mark P, et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space[J]. *Systematic Biology*, 2012, 61(3): 539-542.
- [12] Drummond A J, Suchard M A, Xie D, et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7[J]. *Molecular Biology and Evolution*, 2012, 29(8): 1969-1973.
- [13] Yang J S, Lu B, Chen D F, et al. When did decapods invade hydrothermal vents? Clues from the western Pacific and Indian Oceans[J]. *Molecular Biology and Evolution*, 2012, 30(2): 305-309.
- [14] Zhao Y Y, Zhu X C, Li Y D, et al. Mitochondrial genome of Chinese grass shrimp, *Palaemonetes sinensis* and comparison with other Palaemoninae species[J]. *Scientific Reports*, 2019, 9(1): 17301.
- [15] Shen X, Sun M A. Analysis of mitochondrial genome characteristics and exploration of molecular markers in Palaemonidae[J]. *Fisheries Science*, 2011, 30(6): 347-351. [申欣, 孙名安. 长臂虾科线粒体基因组特征分析及分子标记探讨[J]. *水产科学*, 2011, 30(6): 347-351.]
- [16] Tian M, Shen X, Meng X P, et al. Analysis of mitochondrial genome characteristics and genetic different loci in Brachyuran[J]. *Fisheries Science*, 2011, 30(1): 31-37. [田美, 申欣, 孟学平, 等. 短尾派线粒体基因组特征及基因差异位点分析[J]. *水产科学*, 2011, 30(1): 31-37.]
- [17] Shen X, Sun M A, Wu Z G, et al. The complete mitochondrial genome of the ridgetail white prawn *Exopalaemon carinicauda* Holthuis, 1950 (Crustacean: Decapoda: Palaemonidae) revealed a novel rearrangement of tRNA genes[J]. *Gene*, 2009, 437(1-2): 1-8.
- [18] Shen X, Wang H Q, Wang M X, et al. Comparison of the characteristics and gene order in mitochondrial genomes of krills[J]. *Progress in Fishery Sciences*, 2012, 33(2): 49-55. [申欣, 王海青, 王敏晓, 等. 磷虾类线粒体基因组的特征和基因排列比较[J]. *渔业科学进展*, 2012, 33(2): 49-55.]
- [19] Nei M, Kumar S. *Molecular Evolution and Phylogenetics*[M]. New York: Oxford University Press, 2000.
- [20] Shen X, Li X, Xu Q H. Comparison and analysis of *Alpheus japonicus* and *A. distinguendus* complete mitochondrial genome sequences[J]. *Acta Oceanologica Sinica*, 2012, 34(5): 147-153. [申欣, 李晓, 徐启华. 日本鼓虾与鲜明鼓虾线粒体基因组全序列的分析比较[J]. *海洋学报*, 2012, 34(5): 147-153.]
- [21] Shen X, Ren J F, Cui Z X, et al. The complete mitochondrial genomes of two common shrimps (*Litopenaeus vannamei* and *Fenneropenaeus chinensis*) and their phylogenomic considerations[J]. *Gene*, 2007, 403(1-2): 98-109.
- [22] Zhou Z X, Liu B, Gong J, et al. Phylogeny and population genetics of species in *Takifugu* genus based on mitochondrial genome[J]. *Journal of Fisheries of China*, 2020, 44(11): 1792-1803. [周志雄, 刘波, 宫杰, 等. 基于线粒体基因组的东方鲀属系统发育学及群体遗传学[J]. *水产学报*, 2020, 44(11): 1792-1803.]
- [23] von Nickisch-Rosenegk M, Brown W M, Boore J L. Complete sequence of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: Gene arrangements indicate that Platyhelminths are Eutrochozoans[J]. *Molecular Biology and Evolution*, 2001, 18(5): 721-730.
- [24] Boore J L, Medina M, Rosenberg L A. Complete sequences of the highly rearranged molluscan mitochondrial genomes of the Scaphopod *Graptacme eborea* and the bivalve *Mytilus edulis*[J]. *Molecular Biology and Evolution*, 2004, 21(8): 1492-1503.
- [25] Helfenbein K G, Fourcade H M, Vanjani R G, et al. The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister group to protostomes[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(29): 10639-10643.
- [26] Dellaporta S L, Xu A, Sagasser S, et al. Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(23): 8751-8756.
- [27] Li Y Q, Song J, Shen X, et al. The first mitochondrial genome of *Macrobrachium rosenbergii* from China: Phylogeny and gene rearrangement within Caridea[J]. *Mitochondrial DNA Part B*, 2019, 4(1): 134-136.
- [28] Sung C H, Cheng C C, Lu J K, et al. The complete mitochondrial genome of *Hymenocera picta* (Malacostraca: Decapoda: Hymenoceridae)[J]. *Mitochondrial DNA Part B*, 2018, 3(2): 790-791.
- [29] Bracken H, De Grave S, Felder D L. Phylogeny of the infraorder Caridea based on mitochondrial and nuclear genes (Crustacea: Decapoda)[M]//Decapod Crustacean Phylogenetics. Boca Raton: Taylor & Francis/CRC Press, 2009.
- [30] De Grave S, Fransen C H J M, Page T J. Let's be pals again: Major systematic changes in Palaemonidae (Crustacea: Decapoda)[J]. *PeerJ*, 2015, 3: e1167.

Characteristics and phylogenetic analysis of the mitochondrial genome in Palaemonidae

ZHU Longqiang^{1, 2, 3}, ZHU Zhihuang^{2, 3}, LIN Qi^{2, 3}, ZHU Leiyu^{2, 3}, WANG Dingquan¹, WANG Jianxin¹

1. Laboratory of Molecular Ecology and Application of Marine Microorganisms, Zhejiang Ocean University, Zhoushan 316000, China;
2. Fisheries Research Institute of Fujian Province, Xiamen 361013, China;
3. Key Laboratory of Cultivation and High-Value Utilization of Marine Organisms in Fujian Province, Xiamen 361013, China

Abstract: The basic characteristics of Palaemonidae mitogenomes were revealed by the comprehensive mitogenome analysis of 14 Palaemonidae species. The results showed that the lengths of the 14 mitogenomes varied from 15396 bp to 15967 bp. The content of A + T ranged from 58.97% to 69.09%. Compared with the gene order of a Decapoda ancestor, the gene arrangement order of the *Macrobrachium* had not changed, but the gene arrangement order of *Palaemon*, *Anchistus*, and *Hymenocera* changed to varying degrees. Furthermore, gene deletion was detected in the mitogenome of *Anchistus*. The K_a/K_s ratios of 13 protein coding genes (PCGs) in *Palaemon* and *Macrobrachium* was much lower than one, indicating strong purification. In addition, the codon usage of the 13 PCGs in Palaemonidae showed that the encoded amino acids had similar preferences. Among the two rRNA genes and 13 PCGs identified the mitogenome, the analysis of the different sites between *Palaemon* and *Macrobrachium* showed that the genes *nd5*, *nd4*, and *nd2* were ideal molecular markers. The phylogenetic trees based on the 13 PCGs showed that *Palaemon* and *Macrobrachium* were sisters, and *Anchistus* and *Hymenocera* were gathered together into a large branch. Finally, according to the estimation results of the molecular clock, it is speculated that Palaemonidae species may have originated in the Permian, before further differentiating into a species with modern characteristics. This study provides a reliable molecular marker for the rapid identification of Palaemonidae species and a theoretical basis for analyzing the genetic diversity of Palaemonidae species.

Key words: Palaemonidae; mitogenome; structure characteristic; phylogeny

Corresponding author: WANG Jianxin. E-mail: jxwang@zjou.edu.cn