

DOI: 10.3724/SP.J.1035.2010.00177

鲢中华鳊线粒体基因组的基因组成及特征分析

彭国祥^{1,2} 高 谦^{2,*} 宋 英² 赵琴平³ 罗宇良¹ 聂 品²

(1. 华中农业大学水产学院, 武汉 430070; 2. 中国科学院水生生物研究所, 淡水生态与生物技术国家重点实验室, 武汉 430072; 3. 武汉大学基础医学院, 武汉 430071)

摘要: 研究旨在测定隶属桡足亚纲(Copepoda)杯口水蚤目(Poecilostomatoida)的寄生甲壳动物鲢中华鳊(*Sinergasilus polycolpus*)线粒体 DNA (mtDNA)序列, 分析其结构特征, 为节肢动物系统进化研究积累资料。利用简并引物扩增鲢中华鳊 Cyt b 和 CO I 基因片段, 据此设计特异性引物 Long-PCR 扩增鲢中华鳊 mtDNA 序列, 拼接后得到 12917 bp 的序列(GenBank 登录号为 EU621723)。这段序列 A、T、C、G 碱基含量分别为 35.0%、35.8%、14.0%和 15.2%, GC-Skew 和 AT-Skew 值分别为 0.041 和 -0.011。通过序列同源性和结构分析, 鉴定了 35 个基因, 包括 13 个蛋白编码基因、21 个 tRNA 基因和 1 个 12S rRNA 基因, 推测未能扩增到的片段包含 16S rRNA、tRNA^{Arg} 和 tRNA^{Thr} 基因以及控制区序列。tRNA 基因长度在 54—63bp 之间, 均可折叠成典型的三叶草结构, 但 T ψ C 臂长度变化较大, 其中有三个缺少 T ψ C 臂。35 个基因中, 4 个蛋白质编码基因和 7 个 tRNA 基因位于 N 链, 其余的位于 J 链。基因间重叠有 12 处, 重叠碱基量达 159 bp。进一步研究桡足类及相关类群线粒体基因组的遗传特征将有助于揭示和阐明桡足类系统发育关系。

关键词: 鲢中华鳊; 桡足类; 线粒体基因组; Long-PCR 扩增

中图分类号: Q781 文献标识码: A 文章编号: 1000-3207(2010)01-0177-07

动物线粒体 DNA (Mitochondrial DNA, mtDNA) 为闭合环状双链 DNA 分子, 大小通常在 14—42 kb 之间, 其基因组成高度保守, 一般包括 2 个 rRNA 基因、22 个 tRNA 基因、13 个蛋白编码基因以及非编码的控制区^[1]。由于动物 mtDNA 具有分子量小、含量丰富、进化速率较快、母系遗传等特点, mtDNA 基因序列常用于种群遗传和低级分类阶元系统发育研究^[2, 3]。另一方面, 在动物进化过程中线粒体基因重排属低概率事件, 基因排序在较高分类阶元上仍表现出高的相似性, 一般认为线粒体基因排列方式分析适合于动物较高分类阶元的系统发育关系重建^[4-8]。可见, 动物线粒体 DNA 可在基因排序和序列变异两方面为动物系统进化研究提供有用的信息。

节肢动物种类繁多, 其内部各大类群间的系统发生关系, 至今仍未有明确的结论^[9, 10], 其中桡足亚纲(Copepoda)在甲壳亚门(Crustacea)中的位置及

其内部各目间的关系尚存争议^[11]。随着支序分析方法和分子生物学技术发展, 利用 mtDNA 的基因排列和序列变异信息探讨节肢动物的系统发育已成为研究热点^[2-8]。桡足类包括 200 科、1650 属、11500 种^[12], 但有关桡足类线粒体基因组的研究较少, 迄今仅猛水蚤目的加利福尼亚虎足水蚤(*Tigriopus californicus*)^[13]、日本虎足水蚤(*Tigriopus japonicus*)^[14]和管口水虱目的鲢疮痂鱼虱(*Lepiophtheirus salmonis*)^[15]的完整线粒体基因组被报道。隶属杯口水蚤目(Poecilostomatoida)鳊科(Ergasilidae)的鲢中华鳊(*Sinergasilus polycolpus*)是鲢(*Hypophthalmichthys molitrix*)、鳊(*Aristichthys nobilis*)鳃部寄生的主要寄生虫。宋英等^[17]对鲢中华鳊、大中华鳊(*S. major*)和鲤中华鳊(*S. undulatus*)的系统发育关系及其与鱼类宿主的协同进化进行了研究。本文采用 Long-PCR 扩增鲢中华鳊线粒体基因组序列, 解析其序列结构

收稿日期: 2008-05-15, 修订日期: 2009-02-14

基金项目: 淡水生态与生物技术国家重点实验室自主项目(编号: 2008FBZ06)资助

作者简介: 彭国祥(1982—), 男, 湖北省武汉市人; 硕士; 主要从事鱼类寄生虫学研究。E-mail: guoxiangp@163.com。*同为第一作者

通讯作者: 聂品, E-mail: pinnie@ihb.ac.cn

特征,为深入探讨和阐明桡足类的系统发育奠定基础。

1 材料与方法

1.1 鲢中华鲮总 DNA 提取

实验所用鲢、鳙鱼购自湖北省武汉市汤逊湖渔场,活体运回实验室后,在解剖镜下从鱼鳃上收集鲢中华鲮,用 0.9%生理盐水将虫体漂洗干净,75%酒精固定,−20℃保存。整虫消化后酚-氯仿法抽提总 DNA^[18]。

1.2 线粒体基因组序列的 PCR 及 Long-PCR 扩增

甲壳动物 Cyt *b*、CO I 基因序列的报道较多,根据鲢中华鲮近源物种的 Cyt *b*、CO I 基因序列,设计简并引物,扩增鲢中华鲮 Cyt *b*、CO I 基因片断并进行序列测定。根据已测定序列,设计两对特异性引物,即 spCytb131F: 5'-TAATACCTTTTGTAGTAGC CGCCCTT-3'和 spCOI555R: 5'-CGGTGACGATAAC ATTGTAGATTGG-3', spCOI123F: 5'-CCAAATCTA CAATGTTATCGTCACCGC-3'和 spCytb753R: 5'-CA AATCCTCCCCACAATCAAGTAACAA-3'。25 μL PCR 反应体系包括 2.5 μL 10×PCR 反应缓冲液, dNTPs 100 μmol/L, 正反向引物各 0.15 μmol/L, LA-Taq 聚合酶 0.65 U, 基因组 DNA 50 ng, 灭菌双蒸水补至终体积。PCR 扩增在 MJ Model PTC-100 型 PCR 仪上进行,降落 PCR 反应条件为: 94℃预变性 3min; 94℃变性 10s, 68℃退火延伸 10min, 94℃变性 10s, 67℃退火延伸 10min, 如此梯度降落至 94℃变性 10s, 58℃退火延伸 10min; 然后 94℃变性 10s, 63℃退火 30s, 68℃延伸 10 min, 35 个循环; 68℃延伸 10 min。1.0%琼脂糖凝胶电泳,参照 E.Z.N.A.凝胶纯化试剂盒(Omega)的使用说明,回收目的片段。

1.3 亚克隆的制备、序列测定及基因鉴定

将扩增到的片段用限制性内切酶 *Mbo* I 酶切,分别回收大小在 300—2000 bp 的所有片段,并将其克隆到 pUC19 载体中,筛选阳性克隆测序,用 DNASTAR 软件分析测序结果,并在 NCBI (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>) 中进行比对。以扩增到的大片段为模板,设计引物扩增覆盖缺口的片断并测序,应用 DNASTAR 软件中的 Seqman 程序完成所测线粒体 DNA 序列的连接。

Open Reading Frame Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 软件分析结合 BLAST (<http://ncbi.nlm.nih.gov/BLAST>) 搜索,并与已知同

源蛋白比较,确定并定位鲢中华鲮 13 个蛋白编码基因。用 Mega3.1 软件^[19]分析蛋白编码基因的密码子使用情况。通过近源物种 rRNA 序列比较及二级结构预测,定位鲢中华鲮 rRNA 基因的起始点。利用 tRNAscan-SE1.21 软件(<http://selab.janelia.org/tRNAscan-SE/>)^[20],结合 tRNA 反义密码子和可折叠特性的人工观察,确定鲢中华鲮 tRNA 序列及其结构。

2 结果

2.1 鲢中华鲮线粒体基因组的基因组成与定位分析
拼接后得到 12917 bp 的鲢中华鲮 mtDNA 序列, GenBank 登录号为 EU 621723。尽管未得到完全的线粒体基因组序列,缺少可能的 16S rRNA、tRNA^{Arg} 和 tRNA^{Thr} 基因以及控制区序列,但本研究已鉴定并定位了大部分的鲢中华鲮线粒体基因,包括 13 个蛋白质编码基因、21 个 tRNA 基因和 1 个 12S rRNA 基因(表 1)。

鲢中华鲮线粒体基因组中的蛋白质编码基因包括 7 个 NADH 脱氢酶亚基(NADH dehydrogenase subunit)、2 个 ATP 合成酶亚基(ATP synthase subunit)、3 个细胞色素 c 氧化酶亚基(Cytochrome c oxidase subunit)和 1 个细胞色素 b(Cytochrome b)基因。起始密码子有四种,即 ATA(CO₁、CO₂、ND1、ND2、ND3、ND5 和 ND6)、ATG (ND4L、CO₃ 和 Atp6)、ATC(Cyt *b*)、ATT (Atp8 和 ND4)。四种终止密码子分别是完整的终止密码子 TAA (CO₄、Atp6、Atp8、Cyt *b*、ND1、ND2、ND3、ND4、ND4L 和 ND5)和 TAG (ND6),以及不完整的终止密码子 TA (CO₅)和 T(CO₆)。

在 12917 bp 的鲢中华鲮 mtDNA 序列中,发现 12 处基因重叠,重叠碱基量达到 159 bp; 12 个基因间隔区的总长度为 183 bp (图 1)。在所有已定义的基因中,4 个蛋白质编码基因,即 CO I、ND5、Cyt *b* 和 ND4,以及 tRNA^{Leu(CUN)}、tRNA^{Met}、tRNA^{Pro}、tRNA^{Asp}、tRNA^{Glu}、tRNA^{Tyr} 和 tRNA^{His} 基因位于 N 链,其余基因位于 J 链。

鲎类异形鲎(*Limulus polyphemus*)的线粒体基因排列被认为可代表节肢动物最原始的线粒体基因排列方式^[22],比较鲢中华鲮和异形鲎的线粒体基因排列,发现两者之间仅仅 Atp8-Atp6 和 ND4L-ND6 片段具可比性;与同亚纲的加利福尼亚虎足水

表 1 鲢中华鳊线粒体基因组基因定位特征
Tab. 1 Features of mtDNA sequence of *Sinergasilus polycolpus*

基因名称 Name	起点—终点 Start—End	基因长度 Base-pair	反义密码子 Anti-codon	起始密码子 Start codon	终止密码子 Stop codon	肽链长度 Amino acid
tRNA ^{Ser}	1—57	57	AGA			
<i>Atp 6</i>	35—718	684		ATG	TAA	227
<i>Atp 8</i>	711—794	84		ATT	TAA	27
tRNA ^{Lys}	804—864	61	TTT			
tRNA ^{Leu}	865—927	63	TAG			
tRNA ^{Gly}	928—986	59	TCC			
<i>ND4L</i>	985—1317	333		ATG	TAA	110
tRNA ^{Leu}	1263—1320	58	CAG			
<i>ND6</i>	1310—1786	477		ATA	TAG	158
12S rRNA	1786—2420	635				
tRNA ^{Val}	2421—2478	58	TAC			
<i>CO</i>	2479—3268	790		ATG	T	262
<i>CO</i>	3275—4818	1544		ATA	TA	513
tRNA ^{Met}	4824—4885	62	CAT			
tRNA ^{Pro}	4886—4948	63	TGG			
tRNA ^{Asp}	4946—5002	57	GTC			
tRNA ^{Glu}	5002—5064	63	TTC			
tRNA ^{Ser}	5072—5125	54	TCT			
tRNA ^{Asn}	5170—5229	60	GTT			
<i>ND3</i>	5230—5574	345		ATA	TAA	114
<i>CO</i>	5561—6253	693		ATA	TAA	230
tRNA ^{Trp}	6275—6335	61	TCA			
<i>ND1</i>	6333—7256	924		ATA	TAA	307
tRNA ^{Phe}	7257—7315	59	GAA			
<i>ND5</i>	7320—9023	1704		ATA	TAA	567
tRNA ^{Gln}	9016—9077	62	TTG			
<i>ND2</i>	9078—10031	954		ATA	TAA	317
tRNA ^{Tyr}	10033—10091	59	GTA			
<i>Cyt b</i>	10092—11228	1137		ATC	TAA	378
tRNA ^{Leu}	11222—11282	61	TAA			
<i>ND4</i>	11313—12632	1320		ATT	T	439
tRNA ^{Cys}	12631—12686	56	GCA			
tRNA ^{Ile}	12739—12800	62	GAT			
tRNA ^{Ala}	12800—12856	57	TGC			
tRNA ^{His}	12856—12917	62	GTG			

蚤和鲢疮痂鱼虱的线粒体基因组相比较, 鲢中华鳊线粒体基因排列方式与这两种的也存在显著差异(数据未显示)。

2.2 tRNA 基因的结构特征

所鉴定的 21 个 tRNA 基因中包含 3 个 tRNA^{Leu} 基因, 这些 tRNA 基因的长度在 54—63 bp 之间, 均可折叠成典型的三叶草结构(图 1)。仅 tRNA^{Ser(AGN)}

的 aminoacyl-臂长度为 5 bp, tRNA^{Asp} 和 tRNA^{Leu(CUN)} 的为 6 bp, 其余的均为 7 bp; TψC 臂长度变化较大, tRNA^{Asn}、tRNA^{Gln}、tRNA^{Glu}、tRNA^{Iso}、tRNA^{Leu(CUN)}、tRNA^{Lys} 的为 4 bp, tRNA^{Asp}、tRNA^{His}、tRNA^{LeuUUR}、tRNA^{Met}、tRNA^{Pro} 的为 3 bp, tRNA^{Phe} 和 tRNA^{Tyr} 的为 2 bp, tRNA^{Ala}、tRNA^{Cys}、tRNA^{Gly}、tRNA^{Tyr}、tRNA^{Val} 的为 1 bp, 而 tRNA^{Ser(AGN)}、tRNA^{Ser(UCN)} 和

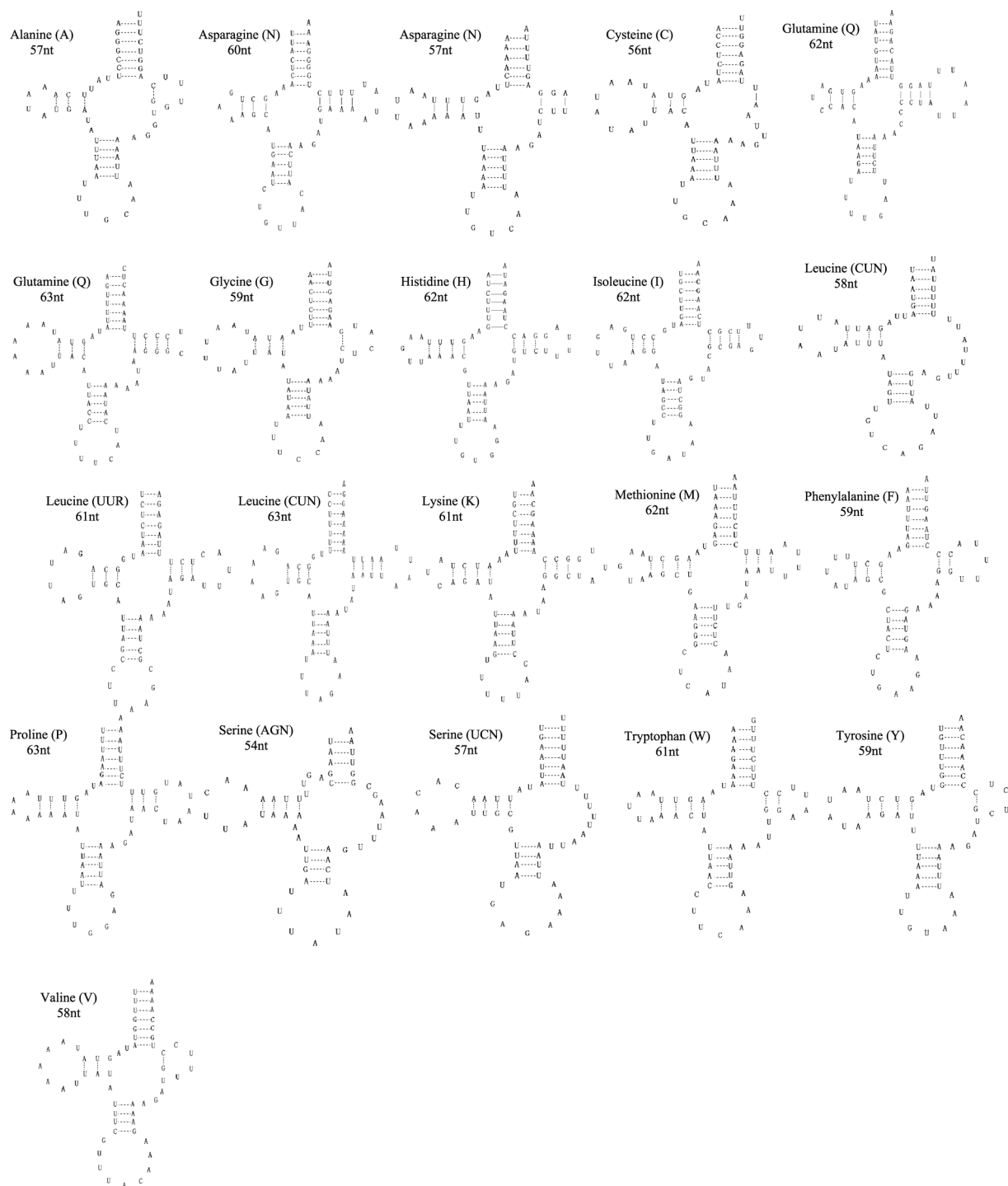


图1 鲢中华鲮 mtDNA 21 种 tRNA 的结构图

Fig. 1 The structure of 21 tRNA genes in *Sinergasilus polycopus* mtDNA

tRNA^{Leu(CUN)}缺少 T ψ C 臂; 这些 tRNA 基因的 DHU 臂长为 3—4 bp。

2.3 碱基组成分析

鲢中华鲮线粒体基因组主编码链四种碱基的含量分别为 A=35.0%、T=35.8%、C=14.0%、G=15.2%，A+T 含量达 70.8% (表 2)。线粒体基因组主编码链

中 G 和 C 含量的不对称性常用 GC-Skew 表示, $GC-Skew = (G-C)/(G+C)$; 类似地, $AT-Skew = (A-T)/(A+T)$ [21]。计算显示, 鲢中华鲮 GC-Skew 和 AT-Skew 值分别为 0.041 和 -0.011。其他的线粒体基因组已报道的鲮足类 [13-16] 的碱基组成及其计算的 GC-Skew 和 AT-Skew 值列于表 2。

表 2 鲢足类线粒体基因组碱基组成特征
Tab. 2 Nucleotide composition of mitochondrial genomes in maxillopods

种类 Taxon	序列长度 Length (bp)	A+T 含量 (A+T) %	GC-不对称 GC-Skew	AT-不对称 AT-Skew
<i>Sinergasilus polycolpus</i>	13154	70.8	0.041	-0.011
<i>Lepiophtheirus salmonis</i>	15445	65.3	0.190	-0.005
<i>Tigriopus japonicus</i>	14628	60.4	0.268	-0.171
<i>Tigriopus Californicus</i>	14539	62.3	0.268	-0.157
<i>Megabalanus volcano</i>	15107	68.3	-0.136	-0.081
<i>Tetracita japonica</i>	15194	66.1	-0.257	0.044
<i>Pollicipes polymerus</i>	15634	67.0	-0.133	0.039
<i>Eucalanus bungii</i>	9530	59.6	-0.003	0.006

3 讨 论

3.1 线粒体基因组扩增问题

本研究获得了 12917 bp 的鲢中华鳊 mtDNA 序列, 但没能获得 *ND4* 和 *Atp6* 基因间可能的 16S rRNA、tRNA^{Arg} 和 tRNA^{Thr} 基因以及控制区序列。Machida, *et al.*^[16]认为真哲水蚤属的 *Eucalanus bungii* 和新哲水蚤属的 *Neocalanus cristatus* 其 mtDNA 控制区较多的茎环结构导致 PCR 扩增和 Sequence Walking 失败。Hickerson, *et al.*^[8]推测长手寄居蟹(*Pagurus longicarpus*) mtDNA 控制区可能含有一些特殊结构导致其线粒体基因组有 300bp 左右没有测通。Le, *et al.*^[23]推测动物 mtDNA 控制区中的重复序列可能导致 *LA-Taq* 聚合酶在 Long-PCR 延伸过程发生跳链。后生动物 16S rRNA 基因中一般存在多茎环结构, 鲢中华鳊 mtDNA 控制区中可能也含有难于扩增的特殊结构, 推测这两个因素的组合可能是鲢中华鳊 *ND4* 和 *Atp6* 基因间序列扩增失败的原因。

3.2 鲢中华鳊 mtDNA 碱基组成特点

在节肢动物中, 甲壳类 mtDNA 的 A+T 含量相对较低, 大多在 70%以下^[6]。然而, 尽管没有包括 16S rRNA 基因和控制区, 鲢中华鳊 mtDNA 的 A+T 含量却达到 70.8%。鲢中华鳊 mtDNA 蛋白编码基因“反 G/C 偏爱(Anti-G/C bias)”的碱基组成特性无疑将影响其蛋白链的氨基酸组成特性, 正因此鲢中华鳊 mtDNA 编码蛋白中 Leu 的总使用频次最高, 达 588 次, 而 Cys 仅出现 32 次。另一方面, 包括鲢中华鳊在内, 线粒体基因组已报道的几种桡足类其

GC-Skew 都为正值, 而 AT-Skew 为负值(表 2), 但目前尚不清楚导致桡足类线粒体主编码链碱基不对称性的确切机制。

3.3 鲢中华鳊 mtDNA 基因排列与定位特征

动物线粒体基因组的一个显著特点是结构紧凑, 部分功能相关的蛋白编码基因相邻接, 并存在碱基重叠, 例如, *ND4L* 和 *ND4* 基因, 以及 *Atp8* 和 *Atp6* 基因之间往往存在 7 bp 左右的碱基重叠^[24]。Boyce, *et al.*^[25]认为基因间的碱基重叠促进线粒体基因组小型化, 而小型线粒体基因组复制所需时间短, 具有一定的自然选择优势。但是, 在无脊椎动物线粒体基因组中 *ND4L* 和 *ND4* 基因, 以及 *Atp8* 和 *Atp6* 基因在位置上相分离并不鲜见^[26]。在鲢中华鳊、加利福尼亚虎足水蚤和日本虎足水蚤这三种桡足类中 *Atp8* 和 *Atp6* 基因都有 8 bp 的碱基重叠, 但 *ND4L* 和 *ND4* 基因并不相邻。另外, 碱基重叠也出现在线粒体蛋白编码基因和 tRNA 基因之间^[2,4,8], 本研究发现了 7 处这类重叠, 特别是 tRNA^{Leu(CUN)} 的基因序列与 *ND4L* 基因的 5'端和 *ND6* 基因的 3'端序列完全重叠。在加利福尼亚虎足水蚤和日本虎足水蚤线粒体基因组中, tRNA^{Trp} 基因序列与 16S rRNA 基因的 5'端序列完全重叠, 针对这一现象, 两种可能的转录机制被提出, 第一种是转录双顺反子分成两部分, 分别形成具有功能的 tRNA^{Trp} 和 16S rRNA; 第二种是 mtDNA 转录产物中的一部分形成具有功能的 tRNA^{Trp}, 另一部分形成 16S rRNA, 但两种基因的产物量相应减少^[13]。上述两种转录模型无疑对探讨鲢中华鳊 mtDNA 中碱基重叠的蛋白编码基因和 tRNA 基因的转录机制也具有指导意义。然而, 动物 mtDNA 中蛋白编码基因间, 以及与 tRNA 基因间碱基重叠的遗传学意义还有待深入研究。

参考文献:

- [1] Crease T J. The complete sequence of the mitochondrial genome of *Daphnia pulex* (Cladocera: Crustacea) [J]. *Gene*, 1999, 233: 89—99
- [2] Munasinghe D H N, Murphy N P, Austin C M. Utility of mitochondrial DNA sequences from four gene regions for systematic studies of Australian freshwater crayfish of the genus *Cherax* (Decapoda: Parastacidae) [J]. *J Crustac Biol*, 2003, 23: 402—417
- [3] Chen H G, Liu W B, Li J Z, *et al.* Comparative analysis of mitochondrial DNA cytb gene and their molecular identifica-

- tion markers in three species of soft shelled turtles [J]. *Acta Hydrobiologica Sinica*, 2006, **30**(4): 380—385 [陈合格, 刘文彬, 李建中, 等. 三种鳖线粒体 DNA 细胞色素 b 基因序列的比较分析. 水生生物学报, 2006, **30**(4): 380—385]
- [4] Yamauchi M, Miya M, Nishida M. Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda) [J]. *Gene*, 2002, **295**: 89—96
- [5] Guo X H, Liu S J, Xiang B, et al. Variability of mitochondrial DNA of triploid hybrids of crucian carp and common carp [J]. *Acta Hydrobiologica Sinica*, 2007, **31**(6): 856—862 [郭新红, 刘少军, 向兵, 等. 三倍体湘云鲫线粒体 DNA 序列变异性分析. 水生生物学报, 2007, **31**(6): 856—862]
- [6] Wilson K, Cahill V, Ballment E, et al. The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods [J]. *Mol Biol Evol*, 2000, **17**: 863—874
- [7] Boore J L, Brown W M. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool [J]. *Curr Opin Genet Dev*, 1998, **8**: 668—674
- [8] Hickerson M J, Cunningham C W. Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, anomura) [J]. *Mol Biol Evol*, 2000, **17**: 639—644
- [9] Giribet G, Edgecombe G D, Wheeler W C. Arthropod phylogeny based on eight molecular loci and morphology [J]. *Nature*, 2001, **413**: 157—161
- [10] Yin W Y. Introduction to a new hypothesis “pancrustacea”: A review of the progress of Arthropod phylogeny [J]. *Zool Res*, 2003, **24**: 11—16 [尹文英. 从泛甲壳动物新假说评述节肢动物系统进化的研究进展. 动物学研究, 2003, **24**: 11—16]
- [11] Huys R, Llewellyn-Hughes J. Small subunit rDNA and Bayesian inference reveal *Pectenophilus ornatus* (Copepoda incertae sedis) as highly transformed Mytilicolidae, and support assignment of Chondracanthidae and Xarifiidae to Lichomolgoidea [J]. *Biol J Linn Soc*, 2006, **87**: 403—425
- [12] Brusca R C, Brusca G J. Invertebrates (Second Edition) [M]. Sunderland Massachusetts: Sinauer Associates, Inc, 2002, 936
- [13] Burton R S, Byrne R J, Rawson P D. Three divergent mitochondrial genomes from California populations of the copepod *Tigriopus californicus* [J]. *Gene*, 2007, **403**: 53—59
- [14] Machida R J, Miya M U, Nishida M. Complete mitochondrial DNA sequence of *Tigriopus japonicus* (Crustacea: Copepoda) [J]. *Mar Biotechnol (NY)*, 2002, **4**: 406—417
- [15] Tjensvoll K, Hodneland K, Nilsen F, et al. Genetic characterization of the mitochondrial DNA from *Lepeophtheirus salmonis* (Crustacea: Copepoda). A new gene organization revealed [J]. *Gene*, 2005, **353**: 218—230
- [16] Machida R J, Miya M U, Nishida M, et al. Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea), with notes on new versatile primers for the srRNA and COI genes [J]. *Gene*, 2004, **332**: 71—78
- [17] Song Y. Phylogenetic Study of Parasitic Copepods in the Ergasilidae and Lernaeidae on Freshwater Fish in China [D]. Doctoral Dissertation, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan. 2006 [宋英. 鳃科和锚头鳃科寄生桡足类的系统发育. 博士学位论文, 中国科学院水生生物研究所, 武汉. 2006]
- [18] Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning: A Laboratory Manual [M]. New York: Cold Spring Harbor Laboratory Press. 2002
- [19] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Brief Bioinform*, 2004, **5**: 150—163
- [20] Lowe T M, Eddy S R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence [J]. *Nucleic Acids Res*, 1997, **25**: 955—964
- [21] Perna N T, Kocher T D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes [J]. *J Mol Evol*, 1995, **41**: 353—358
- [22] Lavrov D V, Boore J L, Brown W M. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus* [J]. *Proc Natl Acad Sci USA*, 2000, **97**: 13738—13742
- [23] Le T H, Blair D, Agatsuma T, et al. Phylogenies inferred from mitochondrial gene orders—a cautionary tale from the parasitic flatworms [J]. *Mol Biol Evol*, 2000, **17**: 1123—1125
- [24] Black W C, Roehrdanz R L. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes [J]. *Mol Biol Evol*, 1998, **15**: 1772—1785
- [25] Boyce T M, Zwick M E, Aquadro C F. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy [J]. *Genetics*, 1989, **123**: 825—836
- [26] Garesse R. *Drosophila melanogaster* mitochondrial DNA: gene organization and evolutionary considerations [J]. *Genetics*, 1988, **118**: 649—663

MITOCHONDRIAL GENES OF *SINERGASILUS POLYCOLPUS* (COPEPODA, ERGASILIDAE) PARASITIZING THE GILLS OF FISH

PENG Guo-Xiang^{1,2}, GAO Qian², SONG Ying², ZHAO Qin-Ping³, LUO Yu-Liang¹ and NIE Pin²

(1. Fisheries College, Huazhong Agricultural University, Wuhan 430070; 2. State key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072; 3. School of Basic Medical Science, Wuhan University, Wuhan 430071)

Abstract: The mitochondrial genomes can provide useful information of the phylogenetic relationships among arthropod groups from two aspects, i.e. gene arrangement and mitochondrial DNA (mtDNA) sequence variation, however, there is no report on the mt genomes of species in the order Poicilostomatoida. The present study aims to determine the nucleotide sequence of mitochondrial genome of *Sinergasilus polycolpus*, a poicilostomatoid copepod, parasitizing the freshwater fish *Hypophthalmichthys molitrix* and *Aristichthys nobilis*. The segments of Cyt *b* and CO I genes of *S. polycolpus* were amplified using the degenerate primers, which were designed on the basis of the homologous sequences. Then the special primers were designed on the above segments to amplify the long mtDNA sequence with Long-PCR. Finally, a mtDNA sequence of 12917 bp in length was gotten from *S. polycolpus*, and deposited in the GenBank database with the accession number of EU621723. This sequence was composed of A of 35.0%, T of 35.8%, C of 14.0% and G of 15.2%, and furthermore the GC-Skew and the AT-Skew were 0.041 and -0.011, respectively. According to the sequence homology and the structure characteristics, the total of 35 genes were identified, including one 12S rRNA, twenty-one tRNA and thirteen protein-coding genes. However, the amplification for the segment containing putative 16S rRNA, tRNA^{Arg} and tRNA^{Thr} genes and control region was unsuccessful. The 13 protein-coding genes had ATA, ATG, ATC and ATT as initiation codons, and TAA, TAG, TA and T as termination codons, respectively. TA and T were incomplete termination codons. All the 21 tRNA genes, ranging between 54 and 63 bp in length, could be folded to form the figure of clover leaf knot. However, other than 3 tRNA genes without T ψ C arm, the length of T ψ C arms of tRNA genes varied from 1bp to 4bp. Of 35 genes, 4 protein-coding genes, i.e. CO I, ND5, Cyt *b* and ND4, and 7 tRNA genes, i.e. tRNA^{Leu(CUN)}, tRNA^{Met}, tRNA^{Pro}, tRNA^{Asp}, tRNA^{Glu}, tRNA^{Tyr} and tRNA^{His}, were encoded on the N strand, the remainders on the J strand. In addition, 12 gene overlapping sites were found in the *Sinergasilus polycolpus* mtDNA, but the significance of gene overlapping in the copepod mt genomes was incompletely clear. In order to elucidate the phylogeny of the copepods, it is necessary to study the mt genomes of more copepods as well as other related groups.

Key words: *Sinergasilus polycolpus*; Copepoda; Mitochondrial genome; Long-PCR