

格氏束腰蟹(*Somanniathelphusa grayi*)的线粒体基因组序列测定和基因顺序进化研究

杜诗雨, 张康琴, 潘 达, 孙红英

(南京师范大学生命科学学院, 江苏省生物多样性与生物技术重点实验室, 江苏 南京 210023)

[摘要] 格氏束腰蟹(*Somanniathelphusa grayi* (Alcock, 1909)) (= *Parathelphusa* (*Parathelphusa*) *chongi* Wu, 1935) 是一种分布于中国云南和印度 Moun Sal 的束腰蟹属淡水蟹物种。本文使用高通量测序技术首次获得了该物种的线粒体基因组近全长序列, 测定的线粒体基因组为一个未闭合的环状 DNA, 长度为 17 654 bp, 包括 37 个基因 (22 个 tRNA 基因, 13 个蛋白质编码基因和 2 个 rRNA 基因), AT 含量为 72.8% (35.1% A, 18.4% C, 8.8% G, 37.7% T)。除 *trnSI-AGN* 缺少一个稳固的双氢尿嘧啶 (DHU) 臂外, 其余 tRNA 基因的二级结构均为典型的三叶草结构。与短尾目线粒体基因组的祖先排列顺序比较发现, 该种线粒体基因组共涉及 6 个 tRNA 基因 (*trnR*, *trnN*, *trnF*, *trnP*, *trnQ* 和 *trnC*) 和 1 个蛋白质编码基因 (*nad5*) 的顺序变化, 与此前报道的两种束腰蟹的基因排列顺序一致。这些基因的重排可以用置换和复制/随机丢失 (TDRL) 模型进行解释。研究发现, 这种重排的线粒体基因组顺序还见于两种在我国江西和海南岛异域分布的束腰蟹种类。提示这种线粒体基因组基因排序是束腰蟹属重要的共近裔特征。本研究可以为格氏束腰蟹的分类厘订, 以及束腰蟹属近缘物种的系统分类和拟地蟹科类群的分子进化研究提供新的分子证据。

[关键词] 格氏束腰蟹, 淡水蟹, 线粒体基因组, 基因重排

[中图分类号] Q959 **[文献标志码]** A **[文章编号]** 1001-4616 (2022) 01-0086-10

The Mitochondrial Genome of *Somanniathelphusa grayi* and the Evolution of Gene Order

Du Shiyu, Zhang Kangqin, Pan Da, Sun Hongying

(School of Life Sciences, Nanjing Normal University, Jiangsu Key Laboratory for Biodiversity and Biotechnology, Nanjing 210023, China)

Abstract: *Somanniathelphusa grayi* (Alcock, 1909) (= *Parathelphusa* (*Parathelphusa*) *chongi* Wu, 1935) was distributed in Yunnan, China and Moun Sal, India. The mitochondrial genome of *S. grayi* was firstly determined using high-throughput sequencing in this study. The sequenced mitochondrial genome of *S. grayi* was nearly-complete, 17 654 bp in length, containing 37 genes (22 tRNA genes, 13 protein-coding genes and 2 rRNA genes). The AT nucleotide content was 72.8% (35.1% for A, 18.4% for C, 8.8% for G and 37.7% for T). The secondary structures of all tRNA genes were typical cloverleaf structures, except for *trnSI-ARN*, which lacked a dihydrouridine (DHU) arm. Compared with the brachyuran ground-pattern mitochondrial gene order, the mitochondrial gene order of *S. grayi* with the positions of six tRNAs (*trnR*, *trnN*, *trnF*, *trnP*, *trnQ* and *trnC*) and one PCGs (*nad5*) have been rearranged, which could be explained by transposition and TDRL models. The gene order was the same as the previously reported two *Somanniathelphusa* species. This gene order was shared by other two *Somanniathelphusa* species from Jiangxi and Hainan Island in China. It is suggested that this gene order pattern is one of the important common characteristics of the genus *Somanniathelphusa*, and it can provide the new molecular evidence for the systematic classification of *Somanniathelphusa grayi*, phylogenetic relationships within *Somanniathelphusa* and the molecular evolution of Gecarcinucidae.

Key words: *Somanniathelphusa grayi*, freshwater crab, mitochondrial genome, gene rearrangement

收稿日期: 2021-02-27.

基金项目: 国家自然科学基金项目 (31772427).

通讯作者: 孙红英, 教授, 研究方向: 动物学、生态学. E-mail: sunhongying@njnu.edu.cn

格氏束腰蟹(*Somanniathelphusa grayi* (Alcock, 1909)) (= *Parathelphusa* (*Parathelphusa*) *chongi* Wu, 1935) 隶属于拟地蟹总科(Gecarcinoidea Rathbun, 1904), 拟地蟹科(Gecarcinidae Rathbun, 1904), 束腰蟹属(*Somanniathelphusa* Bott, 1968), 喜生活在泥洞中^[1-2]。其最显著的特征为雄性腹部第五、六节呈束腰状^[1]。

近年来,高通量测序技术结合生物信息学注释分析的方法已广泛应用于动物线粒体基因组的测定^[3-7]。由于线粒体基因组具有基因组小、大多为母系遗传、进化速率快以及缺少内含子等特点,被广泛地应用于种群遗传、物种鉴定、分子进化和不同分类水平上的系统发生关系等研究^[8]。此外,由于线粒体基因功能上受约束,在DNA序列进化上相对保守,因此在基因的排列和DNA序列进化两个层面上都可以为动物的系统进化研究提供丰富的信息^[9]。

通常情况下,无脊椎动物的线粒体基因组重排情况较为常见,例如头足类^[10]、双壳类^[11]、六足类^[12]和短尾类^[13-14]等。由于近缘物种的线粒体基因顺序相对稳定,因此线粒体基因排列顺序可以用作推断系统发育关系的辅助标记^[15-16]。在已测定线粒体基因组的淡水蟹类中,线粒体基因呈现高度的重排现象,即基因排列顺序均发生了不同程度的重排,共呈现出11种不同于短尾类祖先排列型的排列顺序^[7, 14, 17-20],重排比例高达100%,远高于其他短尾下目49%的物种(目前已知的100种短尾类共有49种发生了线粒体基因重排),但研究的对象主要为溪蟹科淡水蟹类。

迄今为止,已知的拟地蟹科淡水蟹类的线粒体基因组仅束腰蟹属两种:波阳束腰蟹(*Somanniathelphusa boyangensis* Dai, Peng & Zhou, 1994^[17])和坝王束腰蟹(*Somanniathelphusa bawangensis* Dai & Xing, 1994^[14]),这两种线粒体基因组排列呈现出了相同的基因排列顺序^[14]。但是,现有线粒体基因组数据的匮乏,阻碍了拟地蟹科的分子进化和系统分类研究。本文以格氏束腰蟹为研究对象,利用高通量测序技术首次测定了其线粒体基因组近全长序列,并通过序列比对和比较研究,解析了格氏束腰蟹线粒体基因组的结构特征和基因排列顺序。该研究结果为束腰蟹属近缘物种的系统分类和扩散提供了新的分子依据。

1 材料与方法

1.1 样本采集和保存

格氏束腰蟹标本在2015年10月11日采于云南省玉溪市华宁县盘溪镇(103.0631°E, 24.2019°N)。标本在采集并记录生境后立即保存于95%乙醇中。根据《中国动物志》(束腰蟹科,溪蟹科)^[1]在体式显微镜(Nikon SMZ645)下对标本进行分类鉴定。标本保存在南京师范大学生命科学学院,江苏省生物多样性与生物技术重点实验室。

1.2 DNA提取、扩增及质量检测

将存放于95%乙醇中的样本取出,用已灭菌的镊子无菌环境下取适量的鳃组织样本,浸泡于生理盐水中,每30 min更换一次生理盐水,进行脱醇处理。使用动物基因组DNA提取试剂盒(通用型)(Generay Biotech)进行DNA的提取(具体操作步骤参照试剂盒说明书),获得总DNA样品。使用琼脂糖凝胶电泳检测其完整性,并使用微量分光光度计检测其浓度及质量。将质量达标后的DNA样品送至北京诺禾致源科技股份有限公司进行高通量测序。

1.3 序列组装和基因组注释

本文使用MitoZ v1.04^[21]将测序结果(clean data)进行序列组装和注释。随后根据目前NCBI已知的2种束腰蟹(波阳束腰蟹和坝王束腰蟹)线粒体基因组注释结果进行手动校对,从而获得格氏束腰蟹的线粒体基因组全序列。并将注释后的线粒体基因组序列使用MITOS Web Server (<http://mitos.bioinf.uni-leipzig.de/index.py>)进行再次识别确认。

1.4 线粒体基因组组成和结构特征分析

用MEGA v7.0^[22]进行碱基组成、密码子使用情况和蛋白编码基因同义密码子相对使用率(Relative Synonymous Condon Usage, RSCU)的计算。利用通用公式: $AT-skew = (A-T)/(A+T)$, $GC-skew = (G-C)/(G+C)$ 进行碱基偏斜率的计算^[23]。使用在线MITOS Web Server对所有物种的tRNA基因二级结构进行预测。使用VARNA v3-93^[24]和Rnaviz 2.0.3^[25]对预测的tRNA基因二级结构进行绘制,并使用Adobe Photoshop CS 6.0进行最终的修饰与美化。

2 结果与讨论

2.1 线粒体基因组结构特征

格氏束腰蟹线粒体基因组为长度 17 654 bp 的未闭合环状 DNA(图 1),其序列已提交至 GenBank,检索号为 OM214524. 该线粒体基因组共有 37 个基因:13 个蛋白质编码基因(PCGs),22 个转运 RNA 基因(tRNA gene)和 2 个核糖体 RNA 基因(rRNA gene). 由于主非编码区(main non-coding region,*mNCR*)序列有着冗长的 AT 重复,对该区段 DNA 序列的拼接造成干扰. 这种情况在节肢动物线粒体基因组中比较常见. 本文该种的 *mNCR* 所在片段区域在测定时同样遇到此类情况,导致未能被完整测序. 有 23 个基因由正链编码:9 个 PCGs(*atp6*,*atp8*,*cox1*,*cox2*,*cox3*,*cob*,*nad2*,*nad3* 和 *nad6*)和 14 个 tRNA 基因(*trnI*,*trnM*,*trnW*,*trnL2*,*trnD*,*trnK*,*trnG*,*trnA*,*trnR*,*trnN*,*trnS1*,*trnE*,*trnT* 和 *trnS2*);其余 14 个基因由反链编码:4 个 PCGs(*nad1*,*nad4*,*nad4L* 和 *nad5*),8 个 tRNA 基因(*trnQ*,*trnC*,*trnY*,*trnF*,*trnH*,*trnP*,*trnL1* 和 *trnV*)和 2 个 rRNA 基因(*rrnL* 和 *rrnS*)(表 1),这与泛甲壳类和短尾下目线粒体基因组祖先型的编码方向一致^[19,26]. 该线粒体基因组碱基组成 AT 含量高达 72.8%(35.1% A, 18.4% C, 8.8% G, 37.7% T,表 2),其 AT 偏斜率为-0.0357,GC 偏斜率为-0.3529(表 2).

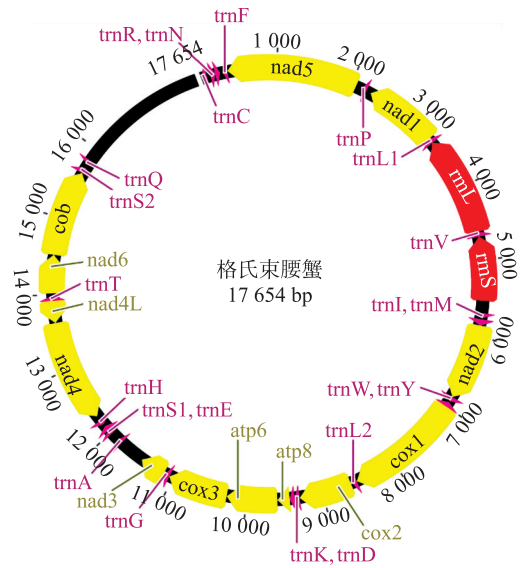


图 1 格氏束腰蟹线粒体基因组结构图
Fig. 1 Mitogenomic map of *Somaniathelphusa grayi*

表 1 格氏束腰蟹线粒体基因组特征及波阳束腰蟹和坝王束腰蟹线粒体基因组特征
Table 1 Mitogenomic features of *Somaniathelphusa grayi*, *S. boyangensis* and *S. bawangensis*

基因	位点			长度/bp			起始密码子/终止密码子			间隔/bp		
	GR	BO	BA	GR	BO	BA	GR	BO	BA	GR	BO	BA
<i>trnC</i> *	18-83	14 755-14 819	10 195-10 259	66	65	65						
<i>trnR</i>	129-196	14 869-14 934	10 312-10 379	68	66	68				45	49	52
<i>trnN</i>	196-265	14 934-15 004	10 379-10 449	70	71	71				-1	-1	-1
<i>trnF</i> *	296-360	15 030-15 092	10 480-10 543	65	63	64				30	25	30
<i>nad5</i> *	360-2 070	15 093-16 803	10 544-12 272	1 711	1 711	1 729	ATA/T	ATA/T	ATG/T	-1	0	0
<i>trnP</i> *	2 144-2 210	16 893-16 958	12 329-12 394	67	66	66				73	89	56
<i>nad1</i> *	2 280-3 251	1-972	12 466-13 425	972	972	960	ATA/TAA	ATA/TAA	ATA/TAA	69	74	71
<i>trnL1</i> *	3 242-3 306	963-1 027	13 428-13 490	65	65	63				-10	-10	2
<i>rrnL</i> *	3 307-4 616	1 028-2 341	13 491-14 800	1 310	1 314	1 310				0	0	0
<i>trnV</i> *	4 617-4 690	2 342-2 414	14 801-14 873	74	73	73				0	0	0
<i>rrnS</i> *	4 691-5 527	2 415-3 248	14 874-15 711	837	834	838				0	0	0
<i>trnI</i>	5 689-5 754	3 437-3 501	15 898-15 962	66	65	65				161	188	186
<i>trnM</i>	5 776-5 843	3 524-3 589	15 985-16 051	68	66	67				21	22	22
<i>nad2</i>	5 843-6 853	3 590-4 600	16 052-17 062	1 011	1 011	1 011	ATG/TAG	ATG/TAG	ATG/TAG	-1	0	0
<i>trnW</i>	6 852-6 917	4 599-4 664	17 061-17 125	66	66	65				-2	-2	-2
<i>trnY</i> *	6 935-6 999	4 684-4 748	17 143-17 208	65	65	66				17	19	17
<i>cox1</i>	7 000-8 535	4 749-6 284	1-1 536	1 536	1 536	1 536	ATG/TAA	ATG/TAA	ATG/TAA	0	0	0
<i>trnL2</i>	8 536-8 599	6 285-6 347	1 537-1 599	64	63	63				0	0	0
<i>cox2</i>	8 619-9 311	6 365-7 048	1 617-2 300	693	684	684	ATG/TAG	ATG/TAA	ATG/TAA	19	17	17
<i>trnK</i>	9 313-9 378	7 059-7 123	2 311-2 375	66	65	65				1	10	10
<i>trnD</i>	9 379-9 442	7 125-7 187	2 377-2 439	64	63	63				0	1	1
<i>atp8</i>	9 442-9 600	7 188-7 346	2 440-2 598	159	159	159	ATG/TAA	ATG/TAA	ATG/TAA	-1	0	0
<i>atp6</i>	9 597-10 268	7 343-8 014	2 595-3 266	672	672	672	ATA/TAA	ATA/TAA	ATA/TAA	-4	-4	-4
<i>cox3</i>	10 268-11 057	8 014-8 803	3 266-4 055	790	790	790	ATG/T	ATG/T	ATG/T	-1	-1	-1
<i>trnG</i>	11 058-11 120	8 804-8 865	4 056-4 117	63	62	62				0	0	0
<i>nad3</i>	11 129-11 473	8 875-9 219	4 118-4 471	345	345	354	ATA/TAA	ATA/TAG	ATC/TAG	8	9	0
<i>trnA</i>	11 818-11 881	9 219-9 280	4 472-4 533	64	62	62				344	-1	0
<i>trnS1</i>	11 989-12 056	9 394-9 459	4 646-4 712	68	66	67				107	113	112
<i>trnE</i>	12 059-12 124	9 462-9 525	4 715-4 781	66	64	67				2	2	2
<i>trnH</i>	12 148-12 214	9 552-9 618	4 809-4 875	67	67	67				23	26	27

续表 1 Table 1 continued

基因	位点			长度/bp			起始密码子/终止密码子			间隔/bp		
	GR	BO	BA	GR	BO	BA	GR	BO	BA	GR	BO	BA
<i>nad4</i> *	12 286-13 623	9 680-11 017	4 938-6 275	1 338	1 338	1 338	ATG/TAG	ATG/TAG	ATG/TAA	71	61	62
<i>nad4L</i> *	13 617-13 919	11 011-11 313	6 269-6 571	303	303	303	ATG/TAA	ATG/TAA	ATG/TAA	-7	-7	-7
<i>trnT</i>	13 922-13 985	11 316-11 378	6 574-6 636	64	63	63				2	2	2
<i>nad6</i>	14 013-14 513	11 407-11 907	6 665-7 165	501	501	501	ATA/TAA	ATA/TAA	ATA/TAA	27	28	28
<i>cob</i>	14 513-15 647	11 907-13 041	7 165-8 299	1 135	1 135	1 135	ATG/T	ATG/T	ATG/T	-1	-1	-1
<i>trnS2</i>	15 648-15 715	13 042-13 108	8 300-8 366	68	67	67				0	0	0
<i>trnQ</i> *	15 772-15 840	13 171-13 238	8 426-8 493	69	68	68				56	62	59

注:GR 为格氏束腰蟹缩写;BO 为波阳束腰蟹缩写;BA 为坝王束腰蟹缩写。* 表示编码在反链上的基因。

在格氏束腰蟹线粒体基因组中,共存在 11 处重叠区,其长度范围在 1 bp~10 bp(表 1);共存在 18 处间隔区,其长度范围在 1 bp~344 bp,其中最长的间隔区位于 *nad3* 和 *trnA* 之间(表 1),而在已公布的波阳束腰蟹和坝王束腰蟹线粒体基因组中,其最长间隔区均位于 *rrnS* 和 *trnI* 之间,长度分别为 188 bp 和 186 bp(表 1)。由于每个线粒体基因组序列具有不等的基因间隔非编码序列(intergenic noncoding sequences,IGNs),在十足目线粒体基因组中也经常发现基因重叠和非编码区^[27-29],淡水蟹即是如此。一些重叠区在淡水蟹类线粒体基因组中是相对保守的,这些重叠区也存在于格氏束腰蟹线粒体基因组中,例如,*atp8* 和 *atp6* 之间的重叠序列为 ATAA;*nad4* 和 *nad4L* 之间的重叠序列为 ATGTTAA^[7,10-11,14,20]。

2.2 蛋白基因

在格氏束腰蟹的 13 个蛋白质编码基因中,所有蛋白质编码基因均为典型的 ATN 起始密码子,其中,*cox1*,*cox2*,*cox3*,*atp8*,*nad2*,*nad4*,*nad4L* 和 *cob* 的起始密码子为 ATG;*atp6*,*nad1*,*nad3*,*nad5* 和 *nad6* 的起始密码子为 ATA。与已有的两种束腰蟹线粒体基因组序列相比,除坝王束腰蟹的 *nad3* 为 ATC,*nad5* 为 ATG 外,其余基因起始密码子均相同(表 1)。而终止密码子中,除 *nad5* 和 *cox3* 的终止密码子为不完整的 TA,*cob* 的终止密码子为不完整的 T 外,其余 10 个蛋白质编码基因均为完整的终止密码子(TAA 或 TAG),和已知的 2 条束腰蟹线粒体情况相同(表 1)。蛋白质编码基因的起始密码子通常是 ATN,但终止密码子主要是 TAA,而 TAG、TA 和 T 很少,这种现象也发生在其他短尾下目物种中^[7,14,17,28-30]。

在 13 个蛋白质编码基因中,*atp8* 的 A+T 含量最高,为 76.9%;*cox1* 的 A+T 含量最低,为 65.2%(表 2)。位于正链上的蛋白质编码基因发生了 T 偏斜(AT 偏斜率为-0.157 0)和 C 偏斜(GC 偏斜率为-0.279 7);而位于反链上的蛋白质编码基因发生了 T 偏斜(AT 偏斜率为-0.199 4)和较明显的 G 偏斜(GC 偏斜率为 0.412 2)(表 2)。

表 2 格氏束腰蟹线粒体基因组碱基组成

Table 2 Nucleotide composition of mitochondrial genome of *Somanniathelphusa grayi*

基因名称	序列长度/bp	碱基组成/%						碱基偏斜率	
		A	C	G	T	A+T	G+C	AT-Skew	GC-Skew
线粒体基因组	17 654	35.1	18.4	8.8	37.7	72.8	27.2	-0.035 7	-0.352 9
蛋白质编码基因	11 133	28.9	15.4	14.5	41.2	70.1	29.9	-0.175 5	-0.030 1
蛋白质编码基因-正链	6 819	29.0	19.9	11.2	39.8	68.8	31.1	-0.157 0	-0.279 7
蛋白质编码基因-反链	4 314	28.9	8.2	19.7	43.3	72.2	27.9	-0.199 4	0.412 2
<i>atp6</i>	672	29.7	21.1	9.4	39.8	69.5	30.5	-0.145 3	-0.383 6
<i>atp8</i>	159	29.5	18.6	4.5	47.4	76.9	23.1	-0.232 8	-0.610 4
<i>cox1</i>	1 536	28.1	19.2	15.3	37.4	65.5	34.5	-0.142 0	-0.113 0
<i>cox2</i>	693	32.6	20.1	10.9	36.4	69.0	31.0	-0.055 1	-0.296 8
<i>cox3</i>	791	27.8	21.4	13.4	37.4	65.2	34.8	-0.147 2	-0.229 9
<i>cob</i>	1 135	29.0	20.5	11.2	39.3	68.3	31.7	-0.150 8	-0.293 4
<i>nad1</i>	972	27.5	8.0	18.9	45.6	73.1	26.9	-0.247 6	0.405 2
<i>nad2</i>	1 011	28.2	20.5	8.0	43.3	71.5	28.5	-0.211 2	-0.438 6
<i>nad3</i>	345	29.2	17.0	11.1	42.7	71.9	28.1	-0.187 8	-0.210 0
<i>nad4</i>	1 338	29.9	8.2	18.9	43.0	72.9	27.1	-0.179 7	0.394 8
<i>nad4L</i>	303	25.0	5.7	24.7	44.7	69.7	30.4	-0.282 6	0.625 0
<i>nad5</i>	1 712	29.6	8.6	19.9	41.9	71.5	28.5	-0.172 0	0.396 5
<i>nad6</i>	501	28.7	18.1	7.2	46.0	74.7	25.3	-0.231 6	-0.430 8
<i>rrnL</i>	1 310	41.4	6.8	17.0	34.8	76.2	23.8	0.086 6	0.428 6
<i>rrnS</i>	837	41.2	6.7	15.9	36.2	77.4	22.6	0.064 6	0.407 1

格氏束腰蟹的密码子使用总数为 3 711 次,其中使用频率最高的密码子是亮氨酸(UUA-Leu)(共被使用 351 次,RSCU=3.53). 蛋白编码基因的同义密码子相对使用率情况如图 2 所示. 其中,亮氨酸 Leu(595 次)、苯丙氨酸 Phe(355 次)和异亮氨酸 Ile(336 次)是使用最频繁的 3 个氨基酸(图 2),约占 34.65%.

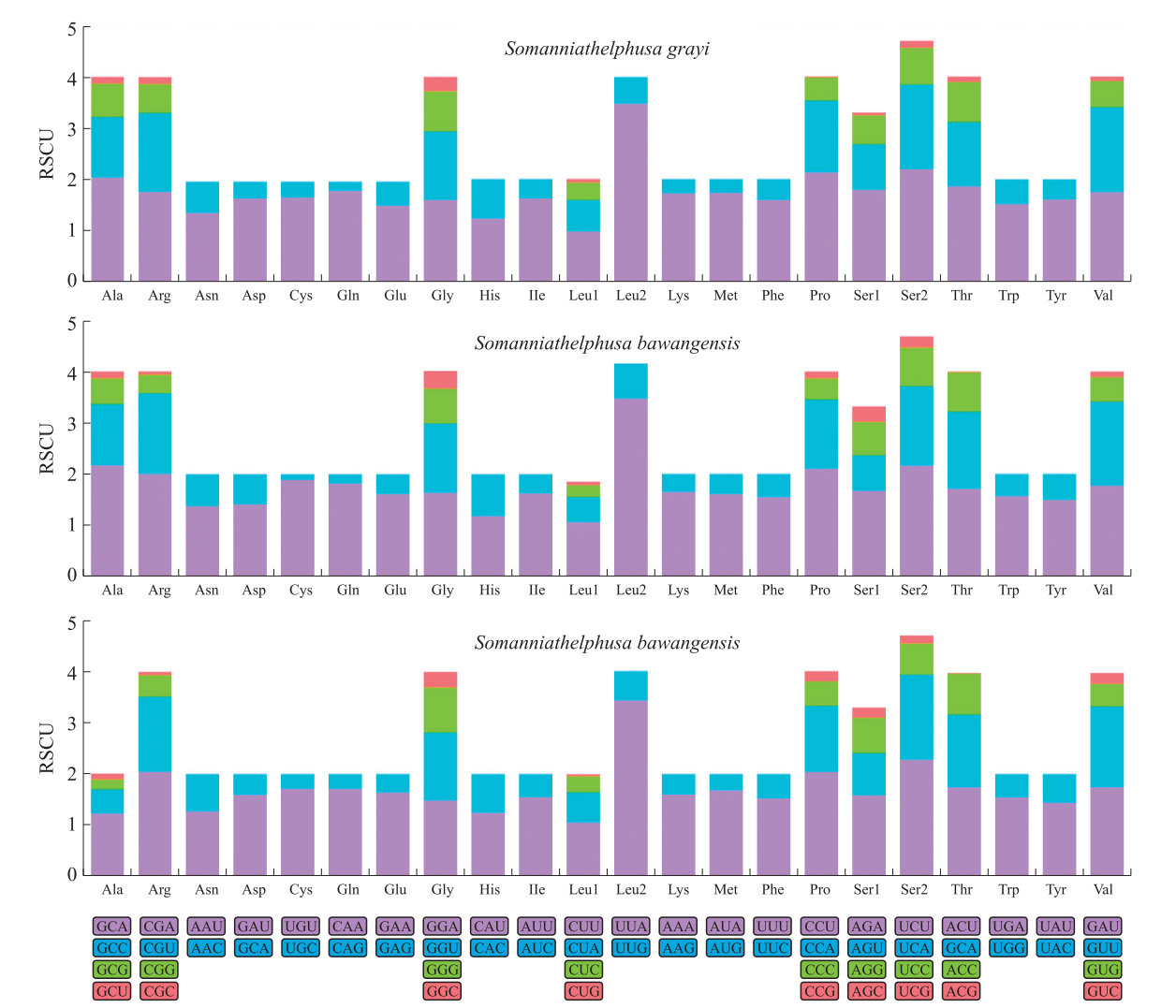


图 2 格氏束腰蟹、波阳束腰蟹和坝王束腰蟹线粒体基因组蛋白质编码基因同义密码子相对使用率
 Fig. 2 The relative synonymous codon usage(RSCU) of 13 PCGs in *Somanniathelphusa grayi*, *S. boyangensis* and *S. bawangensis*

2.3 tRNA 基因和 rRNA 基因

格氏束腰蟹的 22 个 tRNA 基因的二级结构图如图 3 所示. 除了 *trnS1*-AGN 缺少一个稳固的双氢尿嘧啶(DHU)臂,其余所有 tRNA 基因的二级结构均为典型的三叶草结构,即拥有 4 个恒定臂. 其中 *trnS1*-AGN 二级结构 DHU 臂的缺失在大部分后生动物中均有报道^[31-32],且在很多短尾下目物种中也发现了这种情况^[7,17-18,20].

22 个 tRNA 基因的长度范围为 63 bp~74 bp(表 1),由于各可变环长度有所不同,以及 DHU 臂和 TΨC 臂的长度变化使得各 tRNA 基因的长度有所差异^[33]. 其中,15 种 tRNA 基因(*trnA*,*trnR*,*trnD*,*trnC*,*trnQ*,*trnE*,*trnG*,*trnH*,*trnK*,*trnF*,*trnP*,*trnS1*,*trnS2*,*trnY*,*trnV*)共存在 30 对 GU 碱基错配(图 3),多数发生在 DHU 臂,氨基酸接受(AA)臂和反密码子(AC)臂,较少数发生在 TΨC 臂,*trnC* 和 *trnQ* 中的错配数最多,均为 4 对(图 3).

2 个 rRNA 基因,即 *rrnL* 和 *rrnS*,分别位于 *trnL1* 和 *trnV*,*trnV* 和 *trnI* 之间(表 1). *rrnL* 和 *rrnS* 的长度分别为 1 310 bp 和 837 bp,其 AT 含量分别为 76.2%和 77.4%(表 1).

2.4 线粒体基因组重排分析

与短尾下目祖先型线粒体排列顺序(brachyuran ground-pattern mitochondrial gene order, BMGO)相比较

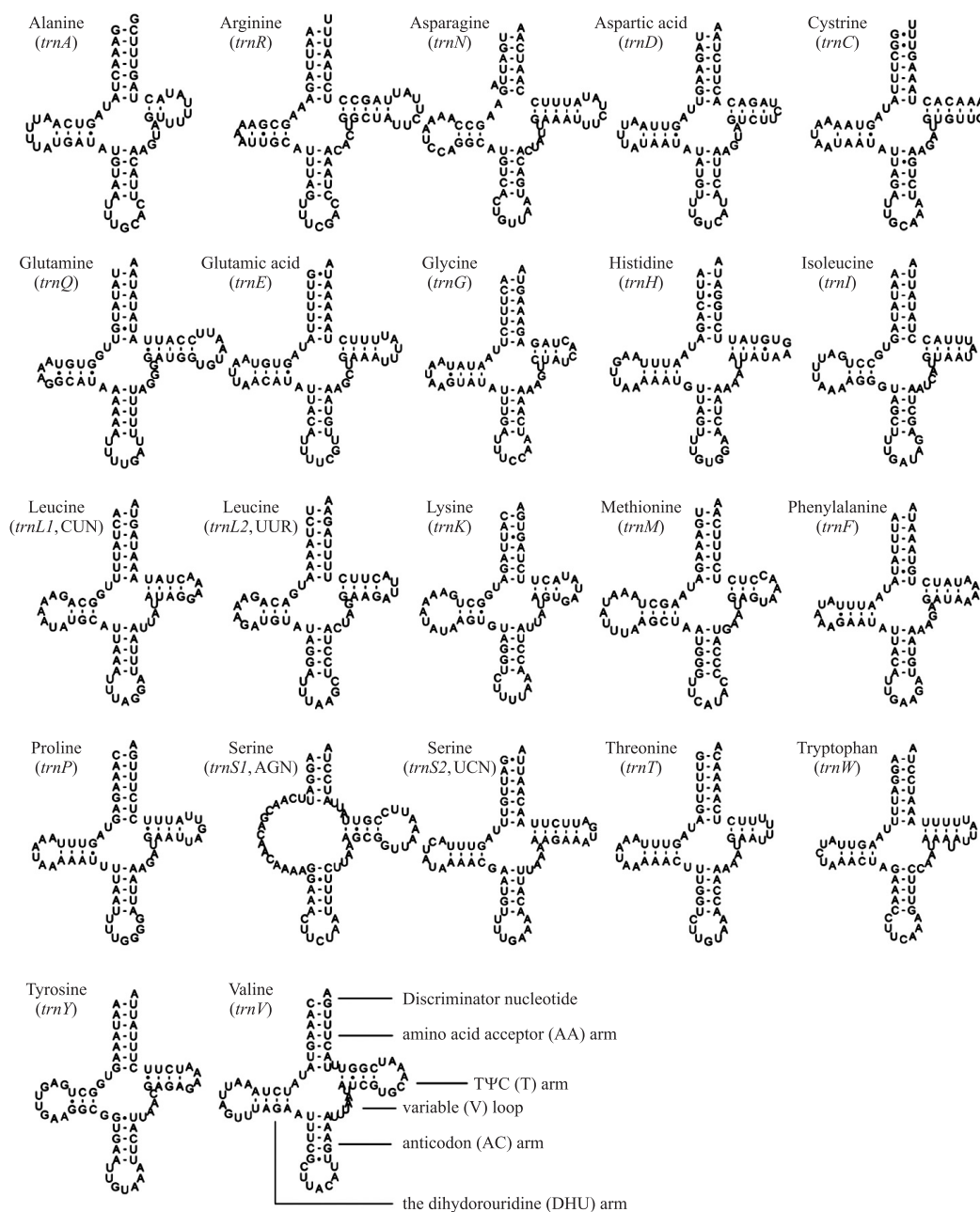


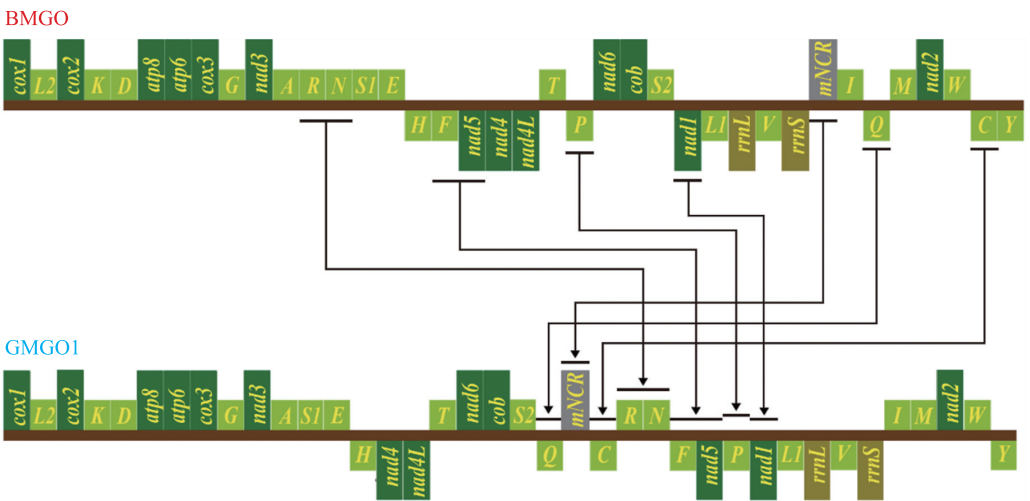
图3 格氏束腰蟹 22 个 tRNA 基因二级结构图

 Fig. 3 Secondary structures of 22 tRNA genes from *Somanniathelphusa grayi*

发现,格氏束腰蟹的线粒体排列顺序发生了较明显的变化. Jia 等^[17]在研究波阳束腰蟹的线粒体基因组排列顺序时,认为 *mNCR* 位于 *rrnS* 与 *trnI* 之间,长度为 188 bp(表 1),AT 含量为 87.8%. 而目前已知的海洋类短尾下目物种的 *mNCR* 长度范围在 332 bp~2 140 bp^[14]. 因此,Zhang 等在分析坝王束腰蟹和波阳束腰蟹的线粒体基因组后,通过序列长度,AT 含量和串联重复元件等综合比较发现,认为坝王束腰蟹和波阳束腰蟹的 *mNCR* 均应位于 *trnC* 与 *trnQ* 之间^[14],修正了波阳束腰蟹的线粒体基因排列顺序,认为波阳束腰蟹与坝王束腰蟹拥有相同的基因排列顺序^[14]. 而本文所研究的格氏束腰蟹线粒体基因组排列顺序与波阳束腰蟹和坝王束腰蟹的线粒体基因排列顺序均相同.

迄今为止,约有 55.2% 的已知的短尾下目物种发生线粒体基因组重排情况. 格氏束腰蟹与此前所报道的 2 种束腰蟹的基因排列顺序一致,即拟地蟹线粒体基因组排列类型 I(gecarcinucid mitochondrial gene order pattern I, GMGO1)(图 4),该重排类型涉及 6 个 tRNA 基因(*trnR*, *trnN*, *trnF*, *trnP*, *trnQ* 和 *trnC*),1 个 PCG(*nad5*)和 1 个 *mNCR* 的排列顺序变化(图 4). 这些基因的重排可以用置换^[34]和复制/随机丢失(TDRL)^[35]模型进行解释. 本文通过以下 3 个假设步骤解释了 GMGO1 的形成过程:首先, *mNCR* 从 [*rrnS*-

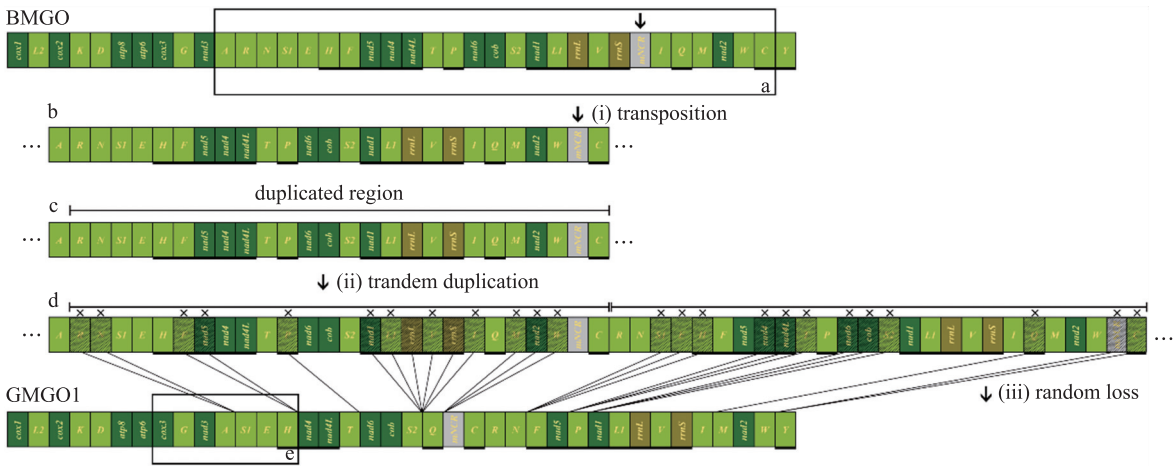
trnI] 基因簇中移出(图 5a), 然后移入[*trnW-trnC*] 基因簇中(图 5b); 其次, 基因簇[*trnR-trnN-trnS1-trnE-trnH-trnF-nad5-nad4-nad4L-trnT-trnP-nad6-cob-trnS2-nad1-trnL1-rrnL-trnV-rrnS-mNCR-trnI-trnQ-trnM-nad2-trnW-trnC*](图 5c) 被串联复制并生成两组相同的基因簇[*trnR-trnN-trnS1-trnE-trnH-trnF-nad5-nad4-nad4L-trnT-trnP-nad6-cob-trnS2-nad1-trnL1-rrnL-trnV-rrnS-mNCR-trnI-trnQ-trnM-nad2-trnW-trnC*](图 5d); 最后, 串联复制后生成的基因簇在随机丢失 26 个基因之后, 形成了目前已知的线粒体基因重排类型, GMGO1.



BMGO:短尾下目祖先型线粒体基因排列顺序;GMGO1:拟地蟹线粒体基因组排列顺序;箭头显示重新排列的基因或基因块

图 4 BMGO 和 GMGO1 两种基因排列顺序的比较

Fig. 4 Comparison of gene arrangements of BMGO and GMGO1



(a) BMGO 基因排列顺序;(b) *mNCR* 的异位;(c) 重复基因区域;(d) 26 个基因的串联重复(16 个 tRNA 基因,7 个 PCGs,2 个 rRNA 基因和 1 和 *mNCR*);*trnR/trnN/trnF/nad5/trnP/nad1/trnL1/rrnL/trnV/rrnS/trnI/trnM/nad2/trnW/trnS1/trnE/trnH/nad4/nad4L/trnT/nad6/cob/trnS2/trnQ/mNCR/trnC*;且复制的 26 个基因随机丢失形成新的基因顺序;(e) GMGO1 基因顺序

图 5 串联重复/随机丢失和重复/反密码子突变假设过程

Fig. 5 The hypothetical process of gene rearrangement in the model of tandem duplication/random loss and duplication/anticodon mutation

2.5 线粒体基因顺序重排的进化含义

对十足目动物的线粒体基因组研究发现,线粒体基因组的排列顺序可能与生态环境转变之间有着重要的联系^[12,36]. 线粒体基因组的排列顺序可能会在适应新环境时发生变化. 例如柯氏绒螯虾(*Shinkaia crosnieri* Baba & Williams, 1998) 和汤花深白蟹(*Gandalfus yunohana*) (Takeda, Hashimoto & Ohta, 2000) 这两类适应极高温环境的十足类,其线粒体基因排列顺序发生了变化^[37-38];一种适应从水生过渡到洞栖陆生生活的一种滑螯虾(*Cherax destructor* (Clark, 1936)),其线粒体排列顺序也发生了变化^[3];中华绒螯蟹

(*Eriocheir japonica sinensis* Milne-Edwards, 1853)的线粒体基因组基因的加速重排可能提示适应非海洋生活过程中的快速辐射进化^[28]. 本文发现3种束腰蟹类的线粒体基因组重排变化较大,共有1个蛋白质编码基因和6个tRNA基因发生重排,这是否与它们对淡水环境的适应性进化有关,还需要通过更多相关类群的线粒体基因组数据进行分子进化研究.

另外,在溪蟹科淡水蟹类发现了线粒体基因组的高度重排现象^[14],但是在拟地蟹科中,是否也会出现与这些已知的重排类型类似或者新的重排方式,这些都还有待进一步扩大拟地蟹科类群的取样,获得必要的线粒体基因组数据,从而挖掘蕴含在线粒体基因组中的进化信息,并探索线粒体基因重排在拟地蟹科类群中的进化含义.

2.6 束腰蟹线粒体基因顺序在其分类和分布的意义

格氏束腰蟹的主要鉴别特征为背部颈沟较为明显,且雄性腹部呈明显的“沙漏形”(hourglass-shaped)^[39]. Alcock^[39]首次描述了格氏束腹蟹(*Parathelphusa* (*Parathelphusa*) *grayi*),模式标本采自印度Moung Sal. 伍献文^[40]描述了常氏束腹蟹(*Parathelphusa* (*Parathelphusa*) *chongi* Wu, 1935),模式标本采自中国云南,该物种系首个由中国学者发现和描述的中国淡水蟹类物种^[1,41]. 随后,这两个物种均被厘定至束腰蟹属(*Somanniathelphusa* Bott, 1968^[42]),即格氏束腰蟹(*Somanniathelphusa grayi* (Alcock, 1909)) and 常氏束腰蟹(*Somanniathelphusa chongi* (Wu, 1935)). 戴爱云等^[43]描述了中华束腰蟹常氏亚种(*Somanniathelphusa sinensis chongi* Dai, Feng, Chen, Song, 1984),模式标本采自中国云南. 随后,戴爱云认为中华束腰蟹常氏亚种(*S. sinensis chongi*)是常氏束腰蟹(*S. changi*)的同物异名^[1],并将该种归并于常氏束腰蟹*S. chongi*. Ng等梳理了世界短尾类分类系统,认为常氏束腹蟹*P. (P.) chongi*(即常氏束腰蟹*S. chongi*)是格氏束腰蟹*S. grayi*的同物异名^[2],并将该种归并于格氏束腰蟹*S. grayi*. 值得注意的是,常氏束腰蟹分布在印度Moung Sai,而格氏束腰蟹*S. grayi*分布在我国云南,地理相隔较远,很可能不是同一物种. 束腰蟹*S. Chongi*属种间雄性腹肢整体形态极为相似,形态特征差异不明显,分类鉴定较为困难,加上缺乏对格氏束腰蟹*S. grayi*正模的详细形态描述^[39],因此常氏束腹蟹*P. (P.) chongi*和格氏束腰蟹*S. grayi*的分类厘定还有待进一步的研究.

线粒体基因排列顺序可以作为特定谱系和分类群的共衍生性状,并为系统发生和进化关系提供了重要的分子证据^[44]. *trnH*基因从*nad5-nad4*基因块易位到*nad3-nad5*之间被认为是短尾下目所共享的特征^[28],本文所研究的格氏束腰蟹和已知的两种束腰蟹均存在*trnH*基因的易位现象^[14,17]. Zhang等在对溪蟹科淡水蟹的线粒体基因排列顺序的进化历程分析时发现,小华溪蟹(*Sinopotamon parvum* Dai, Song, Li, Chen, Wang & Hu, 1985)与小石蟹属(*Tenuilapotamon* Dai, Song, Li, Chen, Wang & Hu, 1984)物种共享一致的重排类型,且亲缘关系较近,认为小华溪蟹的分类地位应重新厘订^[14]. 已知动物线粒体基因组的基因排列顺序相对保守^[44-46],因此不同的类群共享相同或相似的线粒体基因排列顺序可能是来源于共同祖先的结果^[14]. 本研究发现,目前已有线粒体基因组序列的三种束腰蟹分别分布在我国云南(普文、芒掌、思茅、芒洪、景洪、勐洪、小勐养、曼尾、勐罕、勐养、勐海、麻栗坪)、江西(波阳、都昌、沿山)和海南岛(坝王岭、儋县、黎母山、白沙)^[1]. 虽然这3个物种呈异域分布,其分布区之间相隔较远,但线粒体基因排列顺序相一致,提示它们是由共同的祖先进化而来. 对于常氏束腰蟹与格氏束腰蟹的分类关系,线粒体基因组信息(包括DNA序列和基因排列顺序)或可以提供新的证据. 因此,亟需获取分布于印度格氏束腰蟹的线粒体基因组数据做进一步验证.

3 结论

本文首次测定了格氏束腰蟹的线粒体基因组近全长序列,确定其为长度17 654 bp的未闭合环状DNA. 与短尾类祖先类型线粒体基因组比较,格氏束腰蟹线粒体基因组基因顺序与已公布的两种束腰蟹线粒体基因顺序变化一致,共涉及7个基因的重排(1个蛋白质编码基因和6个tRNA基因). 三种在我国云南、江西和海南岛异域分布的束腰蟹种类所共享的线粒体基因排列顺序是该类群重要的共近裔特征之一. 该研究为束腰蟹线粒体基因顺序在其分类和分布方面提供了借鉴,也为拟地蟹科淡水蟹类线粒体基因组信息提供了新的分子证据.

[参考文献]

- [1] 戴爱云. 中国动物志十足目束腹蟹科、溪蟹科[M]. 北京:科学出版社,1999.
- [2] NG P K L, GUINOT D, DAVIE P J F. Systema brachyrotum: part 1. An annotated checklist of extant brachyuran crabs of the world[J]. The raffles bulletin of zoology, 2008, 17: 1-286.
- [3] GAN H M, SCHULTZ M B, AUSTIN C M. Integrated shotgun sequencing and bioinformatics pipeline allows ultra-fast mitogenome recovery and confirms substantial gene rearrangements in Australian freshwater crayfishes[J]. BMC evolutionary biology, 2014, 14(1): 19.
- [4] JEX A R, HALL R S, LITTLEWOOD D T J, et al. An integrated pipeline for next-generation sequencing and annotation of mitochondrial genomes[J]. Nucleic acids research, 2009, 38(2): 522-533.
- [5] TAN M H, GAN H M, SCHULTZ M B, et al. Mitophast, a new automated mitogenomic phylogeny tool in the post-genomic era with a case study of 89 decapod mitogenomes including eight new freshwater crayfish mitogenomes[J]. Molecular phylogenetics and evolution, 2015, 85: 180-188.
- [6] XING Y, MA X, WEI Y, et al. The complete mitochondrial genome of the semiterrestrial crab, *Chiromantes neglectum* (Eubrachyura: Grapsoidea: Sesamidae) [J]. Mitochondrial DNA part B, 2016, 1(1): 461-463.
- [7] XING Y, ZHOU L, HOU Y, et al. Complete mitochondrial genomes from two species of Chinese freshwater crabs of the genus *Sinopotamon* recovered using next-generation sequencing reveal a novel gene order (Brachyura, Potamidae) [J]. ZooKeys, 2017, 705: 41-60.
- [8] BOORE J L. Animal mitochondrial genomes[J]. Nucleic acids research, 1999, 27(8): 1767-1780.
- [9] 孙红英, 周开亚, 宋大祥. 节肢动物线粒体基因组与系统发生重建[J]. 动物学研究, 2003, 24(6): 467-479.
- [10] JIANG L, KANG L, WU C, et al. A comprehensive description and evolutionary analysis of 9 Lologinidae mitochondrial genomes[J]. Hydrobiologia, 2018, 808(1): 115-124.
- [11] WU X, LI X, LI L, et al. New features of Asian *Crassostrea oyster* mitochondrial genomes: a novel alloacceptor tRNA gene recruitment and two novel ORFs[J]. Gene, 2012, 507: 112-118.
- [12] LIU Q N, XIN Z Z, ZHU X Y, et al. A transfer RNA gene rearrangement in the lepidopteran mitochondrial genome [J]. Biochemical and biophysical research communications, 2017, 489(2): 149-154.
- [13] WANG Z, WANG Z, SHI X, et al. Complete mitochondrial genome of *Parasesarma affine* (Brachyura: Sesamidae): gene rearrangements in Sesamidae and phylogenetic analysis of the Brachyura [J]. International journal of biological macromolecules, 2018, 118: 31-40.
- [14] ZHANG Z, XING Y, CHENG J, et al. Phylogenetic implications of mitogenome rearrangements in East Asian potamiscine freshwater crabs (Brachyura: Potamidae) [J]. Molecular phylogenetics and evolution, 2020, 143: 106669.
- [15] BOORE J L, LAVROV D V, BROWN W M. Gene translocation links insects and crustaceans [J]. Nature, 1998, 392(6677): 667.
- [16] LAVROV D V, LANG B F. Poriferan mtDNA and animal phylogeny based on mitochondrial gene arrangements [J]. Systematic biology, 2005, 54(4): 651-659.
- [17] JIA X, XU S, BAI J, et al. The complete mitochondrial genome of *Somanniathelphusa boyangensis* and phylogenetic analysis of Genus *Somanniathelphusa* (Crustacea: Decapoda: Parathelphusidae) [J]. PLoS one, 2018, 13: e0192601.
- [18] SEGAWA R D, AOTSUKA T. The mitochondrial genome of the Japanese freshwater crab, *Geothelphusa dehaani* (Crustacea: Brachyura): evidence for its evolution via gene duplication [J]. Gene, 2005, 355: 28-39.
- [19] TAN M H, GAN H M, LEE Y P, et al. ORDER within the chaos: insights into phylogenetic relationships within the Anomura (Crustacea: Decapoda) from mitochondrial sequences and gene order rearrangements [J]. Molecular phylogenetics and evolution, 2018, 127: 320-331.
- [20] WANG Y, ZENG Q, ZOU J, et al. The complete mitochondrial genome of freshwater crab *Sinopotamon xiushuiense* (Decapoda: Brachyura: Potamoidea) [J]. Mitochondrial DNA part B, 2016, 1(1): 750-752.
- [21] MENG G, LI Y, YANG C, et al. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization [J]. Nucleic acids research, 2019, 47(11): e63.
- [22] KUMAR S, STECHER G, TAMURA K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. Molecular biology and evolution, 2016, 33(7): 1870-1874.
- [23] PERNA N T, KOCHER T D. Unequal base frequencies and the estimation of substitution rates [J]. Molecular biology and

- evolution, 1995, 12(2): 359–361.
- [24] DARTY K, DENISE A, PONTY Y. VARNA: interactive drawing and editing of the RNA secondary structure[J]. Bioinformatics, 2009, 25(15): 1974–1975.
- [25] DE RIJK P, WUYTS J, WACHTER R D. RnaViz 2: an improved representation of RNA secondary structure[J]. Bioinformatics, 2003, 19(2): 299–300.
- [26] BASSO A, BABBUCCI M, PAULETTO M, et al. The highly rearranged mitochondrial genomes of the crabs *Maja crispata* and *Maja squinado* (Majidae) and gene order evolution in Brachyura[J]. Scientific reports, 2017, 7(1): 4096.
- [27] LIU Y, CUI Z. Complete mitochondrial genome of the Asian paddle crab *Charybdis japonica* (crustacea: decapoda: portunidae): gene rearrangement of the marine brachyurans and phylogenetic considerations of the decapods[J]. Molecular biology reports, 2010, 37: 2559–2569.
- [28] SUN H, ZHOU K, SONG D. Mitochondrial genome of the Chinese mitten crab *Eriocheir japonica sinensis* (brachyura: thoracotremata: grapsoidae) reveals a novel gene order and two target regions of gene rearrangements[J]. Gene, 2005, 349: 207–217.
- [29] YAMAUCHI M M, MIYA M U, NISHIDA M. Complete mitochondrial DNA sequence of the swimming crab, *Portunus trituberculatus* (crustacea: decapoda: brachyura) [J]. Gene, 2003, 311: 129–135.
- [30] SHI G, CUI Z, HUI M, et al. The complete mitochondrial genomes of *Umalia orientalis* and *Lyreidus brevifrons*: the phylogenetic position of the family Raninidae within Brachyuran crabs[J]. Marine genomics, 2015, 21: 53–61.
- [31] STOROZHENKO S. To the knowledge of the tribe *Melanoplina* (Orthoptera: Acrididae: Catantopinae) of the Eastern Palearctica [J]. Articulata, 1993, 8(2): 1–22.
- [32] SAHYOUN A H, HÖLZER M, JÜHLING F, et al. Towards a comprehensive picture of alloacceptor tRNA remodeling in metazoan mitochondrial genomes[J]. Nucleic acids research, 2015, 43: 8044–8056.
- [33] CLARY D O, WOLSTENHOLME D R. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code[J]. Journal of molecular evolution, 1985, 22: 252–271.
- [34] MACEY J R, LARSON A, ANANJEVA N B, et al. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome[J]. Molecular phylogenetics and evolution, 1997, 14: 91–104.
- [35] MORITZ C, BROWN W M. Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards[J]. PNAS, 1987, 84: 7183–7187.
- [36] NAKAJIMA Y, SHINZATO C, KHALTURENA M, et al. The mitochondrial genome sequence of a deep-sea, hydrothermal vent limpet, *Lepetodrilus nux*, presents a novel vetigastropod gene arrangement[J]. Marine genomics, 2016, 28: 121–126.
- [37] YANG J S, YANG W J. The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): a novel arrangement and incomplete tRNA suite[J]. BMC genomics, 2008, 9: 1.
- [38] YANG J S, NAGASAWA H, FUJIWARA Y, et al. The complete mitogenome of the hydrothermal vent crab *Gandalfus yunohana* (crustacea: decapoda: brachyura): a link between the Bythograeioidea and Xanthoidea[J]. Zoologica scripta, 2010, 39: 621–630.
- [39] ALCOCK A. Diagnoses of new species and varieties of freshwater crabs[J]. Records of the Indian Museum, 1909, 4: 375–376.28.
- [40] WU H. On a new river crab, *Parathelphusa* (*Parathelphusa*) *chongi*, sp. nov. [J]. Chinese journal of zoology, 1935, 1: 69–73.
- [41] 沈嘉瑞, 戴爱云. 中国动物图谱(甲壳动物 第二册) [M]. 北京: 科学出版社, 1964.
- [42] BOTT R. Parathelphusiden aus hinterindien (crustacea, decapoda, parathelphusidae) [J]. Senckenbergian biologica, 1968, 49(5): 403–422.
- [43] 戴爱云, 冯钟琪, 陈国孝, 等. 中国医学甲壳动物 [M]. 北京: 科学出版社, 1984.
- [44] LI H, LIU H, SHI A, et al. The complete mitochondrial genome and novel gene rearrangement of the unique-headed bug *Stenopirates* sp. (hemiptera: enicocephalidae) [J]. PloS one, 2012, 7(2): e29419.
- [45] BOORE J L. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals [M]. Germany: Springer Netherlands, 2000.
- [46] BLACK W C, ROEHRDANZ R L. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes[J]. Molecular biology and evolution, 1998, 15: 1772–1785.

[责任编辑: 黄 敏]