环境昆虫学报 2021,43 (1):130-146 Journal of Environmental Entomology



王月然,叶飞,门宇,王艳会,谢强.大头隆胸长蝽线粒体基因组测序及分析(半翅目:地长蝽科) [J].环境昆虫学报,2021,43(1): 130-146.

大头隆胸长蝽线粒体基因组测序及分析 (半翅目: 地长蝽科)

王月然¹²³, 叶 飞²³, 门 宇²³, 王艳会²³, 谢 强^{23*}

(1. 南开大学生命科学学院,天津 300071; 2. 中山大学有害生物控制与资源利用国家重点实验室,广州 510275;3. 中山大学生命科学学院,广州 510275)

摘要: 地长蝽科隶属于半翅目异翅亚目蝽次目长蝽总科, 该科包括 15 个族, 缢胸族是其中包含属最多的族, 而目 前该族物种尚无线粒体基因组报道。该族中的隆胸长蝽属昆虫是一个仅在东洋界分布的类群,其中大头隆胸长蝽 Eucosmetus incisus (Walker, 1872) 在中国水稻种植地区广泛发生,是重要的水稻害虫。因此,对大头隆胸长蝽开 展线粒体基因组研究具有重要意义。本研究测定了大头隆胸长蝽线粒体基因组编码区域的全部基因序列,并分析 了其主要特征,结果如下:(1)共测得大头隆胸长蝽线粒体基因序列长度为14562bp,由一部分控制区和典型的 37 个基因组成,包括 22 个转运 RNA 基因,13 个蛋白编码基因和 2 个核糖体 RNA 基因。其线粒体基因排列顺序 同果蝇 Drosophila yakuba 和大多数蝽次目昆虫排列顺序相同。(2) 除 tRNA-His 缺少 TuC 环、不能正常折叠外,其 它 21 个 tRNA 均能折叠成经典三叶草结构。16S rRNA 的结构域 IV 和 V 比结构域 I、II、VI 更保守, 12S rRNA 的 结构域 III 比结构域 I 和 II 更保守。在蝽次目昆虫中,存在两个较稳定的重叠区域,分别位于 ATP8 和 ATP6, ND4 和 ND4L 之间,并且这两段基因重叠区互为反向互补序列(ATGATAA)。(3)核苷酸组成和密码子的使用都表现 出了很高的 AT 偏向性,在13个蛋白编码基因和2个核糖体 RNA中,由N链编码的基因都是 TA 偏移和 GC 偏移, 而除 COI 之外所有由 J 链编码的基因都刚好相反,都是 AT 偏移和 CG 偏移,COI 为 TA 偏移和 CG 偏移。使用最频 繁的密码子均由 AT 组成,且多数不与 tRNA 反密码子严格配对。本文为对缢胸族昆虫线粒体基因组序列的首次报 道,为将来开展缢胸族昆虫相关的分子系统发育研究初步提供了基础数据。除对大头隆胸长蝽本身的分析外,我 们还联合了蝽次目毛点类中其它物种的同源序列,针对红蝽总科、缘蝽总科和长蝽总科间的系统发育关系进行了 研究,结果表明了长蝽总科的单系性,与其亲缘关系最近的是缘蝽总科。 关键词:半翅目;地长蝽科;缢胸族;线粒体基因;系统发育 中图分类号: Q965; S433 文章编号: 1674-0858 (2021) 01-0130-17 文献标识码: A

Sequencing and analyses of the mitochondrial genome of *Eucosmetus incisus* (Hemiptera: Rhyparochromidae)

WANG Yue-Ran^{1,2,3}, YE Fei^{2,3}, MEN Yu^{2,3}, WANG Yan-Hui^{2,3}, XIE Qiang^{2,3*} (1. Nankai University, College of Life Sciences, Tianjin 300071, China; 2. Sun Yat-sen University, State Key Laboratory of Biocontrol, Guangzhou 510275, China; 3. Sun Yat-sen University, School of Life Sciences, Guangzhou 510275, China)

Abstract: Rhyparochromidae belongs to Heteroptera of Hemiptera which includes 15 tribes. Among them,

基金项目: 国家自然科学基金面上项目 (31572242)

作者简介: 王月然,女,硕士,主要研究方向为地长蝽科昆虫系统发育,E-mail: yueranw@mail.nankai.edu.cn

^{*}通讯作者 Author for corresponding: 谢强,博士,教授,主要研究方向为昆虫系统发育,E-mail: xieq8@mail.sysu.edu.cn 收稿日期 Received: 2019-12-25;接受日期 Accepted: 2020-02-12

Myodochini is the biggest tribe which contains many genera, but no mt-genome (mitochondrial genome) has been reported. The genus Eucosmetus is geographically restricted in the Oriental Region. Eucosmetus incisus, an important rice pest occurring widely in China, makes a serious impact on rice yield. Therefore, the molecular phylogenetic research on E. incisus has a great significance. This study determined the gene sequence of the coding region of E. incisus and analyzed its main characteristics of mt-genome. The results were as follows: (1) The part of mt-genome that had been sequenced was 14 562 bp in length and consists of a portion of the control region and 37 typical genes, including 13 protein coding genes (PCGs), 22 transfer RNA (tRNA) genes and 2 ribosomal RNA (rRNA) genes. The gene arrangement was similar to that of Drosophila yakuba and most insects of Pentatomomorpha. (2) In E. incisus, tRNA-His which lacks the T4C loop, could not fold normally, while the other 21 tRNAs could fold into a classical cloverleaf structures. Domains IV and V of the 16S rRNA were more conserved than domains I, II and VI, and domain III of the 12S rRNA was more conserved than domains I and II. In Pentatomomorpha insects , two stable overlapping regions, which located between ATP8 and ATP6, ND4 and ND4L, were reverse complementary sequences (ATGATAA) . (3) The nucleotide composition and codon usage were significantly biased toward A/T. The nucleotide skew statistics for the mt-genome of E. incisus revealed that the J-strand PCGs except COI were AT-skewed and CG-skewed, COI was TA-skewed and CG-skewed, whereas the N-strand PCGs are GC-skewed and TA-skewed. The most prevalent codons were all composed of A/T, and the most codons were not strictly matched with tRNA anticodons. As the first representative from the tribe Myodochini, the mt-genome of E. incisus provided the basic data for future phylogenetic study on insects of Myodochini. In addition to its analysis, we also combined the known sequences of Trichophora and carried out phylogenetic studies of Lygaeoidea, Coreoidea, and Pyrrhocoroidea. Phylogenetic analysed well confirmed the monophyly of Lygaeoidea and provided strong support that Coreoidea was the sister group to the Lygaeoidea.

Key words: Hemiptera; rhyparochromidae; myodochini; mitochondrial genome; phylogenetic analyses

大头隆胸长蝽 Eucosmetus incisus (Walker, 1872) 隶属于半翅目 Hemiptera 异翅亚目 Heteroptera 蝽次目 Pentatomomorpha 长蝽总科 Lygaeoidea 地长蝽科 Rhyparochromidae 缢胸族 Myodochini 隆胸长蝽属 Eucosmetus。缢胸族是地长 蝽科中包含属最多的族。然而,目前尚无缢胸族 相关的线粒体基因组报道。

因此无论是基于其地理分布特殊性还是害虫

防治,都有必要对大头隆胸长蝽开展线粒体基因 组研究,这也可以为地长蝽科昆虫的分子系统发 育研究积累更多数据。本研究扩增并测定了大头 隆胸长蝽线粒体基因组编码区域的全部基因序列, 并且分析了该线粒体基因组的核苷酸组成、密码 子使用、组成偏差、rRNA 二级结构等特征。除对 大头隆胸长蝽线粒体基因组本身的分析外,还联 合了蝽次目毛点类(包括蝽总科、红蝽总科、缘 蝽总科和长蝽总科)(Tullgren,1918; Schaefer, 1975; Schuh and Slater,1995)中其它物种的同源 序列,基于13种蛋白质编码基因构建系统发育 树,对红蝽总科、缘蝽总科和长蝽总科间的系统 发育关系进行了探讨。

1 材料与方法

1.1 标本采集

大头隆胸长蝽的成虫个体,在2018年7月采 自于广西来宾市金秀县青山瀑布,标本浸泡在无 水乙醇中储存,带回实验室后置于 - 20℃环境中 保存。

1.2 DNA 提取

通过解剖获得大头隆胸长蝽胸部肌肉组织, 并用液氮迅速冷冻研磨,利用 CTAB 法(Reineke *et al.*, 2010) 结合试剂盒法提取全基因组, -20℃环境中保存。

1.3 引物设计、PCR 扩增和测序

利用前期设计并优化的通用引物扩增若干具 有部分重叠区域的 DNA 片段(Li *et al.*, 2013), 扩增失败的 DNA 片段,根据已扩增的上下游片段 的序列,利用 Primer premier 5(Lalitha *et al.*, 2000) 设计特异性引物,得到互相重叠的线粒体 基因的片段(表1)。使用 TaKaRa LA DNA 聚合酶 进行 PCR 扩增,PCR 反应过程:94℃ 预变性 2 min,94℃变性 30 s,引物最低扩增温度下退火 45 s,72℃延伸1~3 min,共进行 35 个循环, 72℃终延伸7 min。使用1%琼脂糖凝胶电泳检测 PCR 产物,若条带单一并符合预期大小,则送交 测序服务公司测序(BGI,广州; TSINGKE,广 州)。测序浓度没有达到要求的可通过增加模板量 或者凝胶回收进行处理以提高浓度再测序。

	10,010 1 11111	
片段 Fragment	引物名称 Primer name	引物序列 Nucleotide sequence (5´-3´)
1	TIF	AAAGGRHTATYTTGATA
	TWR	TTAACTITGAAGGYTAATAGTTT
2	ND2F	GATATTARATTGCAAATYTAAAG
	1300R	CTTTARATITGCAATYTAATATC
3	COI-1 F	GGTCAACAAATCATAAAGATATTGG
	COI-2R	GCTCGDGTRTCTACRTCTATWCC
4	C1C3F	GCTCACTTTCATTATGTCTTATCTA
	A8C3R	CTCAAAAAAAGAAATGAAAAA
5	C3N3F2	TCTGGAATTTCCATTACATG
	N3TRR1	ATTGAAGATATTAAGATAGA
6	TRTFF1	TTTACAATTAGTTTCGACCT
	TFTER1	TTTAAGTCAGTTAGGTATGA
7	N5-EF1	GGCAACCCACATAATGATAA
	N5-ER1	CTATTTGATTGAATGTCCCT
8	N5F	AAAGGAATYTGAGCWCTYYTAGT
	N4R	AAYTCWWTYATTGCTTATTCTTC
9	N4LF	AAAAGTAGGAACAATTCTAG
	N6CBR	ATGTAATGTAAAAAATCGTGT
10	CB-10933F	GTWYTWCCNTGAGGRCAAATRTC
	CB-11526R	TCWGCDGGNCGDGCHCCAATTCA
11	CB-11335F	CAYATTCAACCHGAATGRTA
	N1R	GTWGGDTTTGTWACTTTAHTDGARCG
12	N116SF	AACCAGATAAAGTAACCCCA
	N116SR	GACGAGAAGACCCTATAGAA
13	1612SF	TCACTGAGCAGAATAGACCT
	1612SR	TTAATTGATAATCCACGTTG
14	CRF	ΑΤΑΑΤCTAAACCAAACTAAA
	CB-200	TTCTATANDYAGGGTATGAACC

表 1 本研究中使用的 PCR 引物 Table 1 Primers used in this study

1.4 DNA 序列拼接、注释及分析

通过在 GenBank 中进行 BLAST 搜索以初步判断所得序列属于目标类群,用 DNA star v7.1 中的 SeqMan (Swindell and Plasterer, 1997) 组件对序列 进行拼接,得到大头隆胸长蝽的线粒体基因组 序列。

tRNA 基因通过 Mitos WebServer (http: // mitos. bioinf. uni-leipzig. de/index. py) (Bernt *et al.*, 2013) 和 tRNA scan-SE v2.0 (http: //lowelab. ucsc. edu/tRNAscan-SE/) (Lowe and Chan, 2016) 这两种在线工具进行确认并预测二级结构。对于 两种方法都无法检测到的 tRNA 基因,将其和已知 亲缘关系较近的昆虫线粒体基因组的 tRNA 序列进 行比对,确定位置和序列,再利用 RNA structure v5.8 (Mathews and Reuter, 2010) 推测二级结构。

蛋白质编码基因利用 NCBI 中的开放阅读框查 找工具 ORF Finder (https: //www.ncbi.nlm.nih. gov/orffinder/)进行初步注释,再将所得蛋白质编 码基因序列与近缘物种的线粒体基因组相应序列 进行比对,进一步确定各个蛋白质编码基因的边 界,同时获取起始密码子及终止密码子。

核糖体 RNA 基因通常被认为处于两侧基因间 的空缺区域(Boore, 2001; Cameron, 2014b),如 tRNA-Leu (UAG) 和 tRNA-Val 之间的序列界定为 16S rRNA, tRNA-Val 和控制区之间的序列界定为 12S rRNA。核糖体 RNA 的二级结构参考黑腹果蝇 Drosophila melanogaster (双翅目: 果蝇科) (Cannone et al., 2002), 白斑地长蝽 Panaorus albomaculatus (半翅目: 地长蝽科) (Li et al., 2016a) 以及豆突眼长蝽 Chauliops fallax (半翅目: 束长蝽科) (Li et al., 2013) 的二级结构进行预 测, 茎环结构的命名参考烟草天蛾 Manduca sexta (鳞翅目: 天蛾科) (Cameron and Whiting, 2008) 和意大利蜜蜂 Apis mellifera (膜翅目: 蜜蜂科) (Gillespie et al., 2006) 的线粒体 rRNA 基因。确 定线粒体基因组内各个基因的位置后使用 CGView Server (http: //stothard. afns. ualberta. ca/cgview _ server/index. html) (Grant and Stothard, 2008) 在 线平台绘制线粒体基因组的结构示意图。

线粒体基因组的核苷酸组成和蛋白质编码基因的密码子使用频率通过 MEGA v7.0 (Kumar *et al.*,2016) 进行分析; 计算各基因的核苷酸组成偏向性: AT – skew = (A – T) / (A + T)和 GC – skew = (G – C) / (G + C) (Perna and Kocher, 1995)。

1.5 系统发育分析

选取蝽次目内 21 个物种的蛋白质编码基因序 列分析长蝽总科的系统发育地位,其中长蝽总科 的 10 个物种作为内群,蝽总科、红蝽总科、缘蝽 总科的 11 个物种作为外群(表 2)。除大头隆胸长 蝽外的其它物种线粒体基因组均来自 GenBank,按 照物种名进行排序形成了 13 个蛋白质基因的数据 集合,在 MEGA v7.0 中进行多重比对并手工校正, 得到包含 13 个蛋白质编码基因全部三个位点的矩 阵 PCG123。使用矩阵 PCG123 通过 MEGA v7.0 (Kumar et al., 2016) 获得去掉蛋白质编码基因第 3 位密码子的矩阵 PCG12。

本研究采用贝叶斯分析 (Bayesian inference) 和最大似然法 (maximum likelihood) 重建系统发 育树,核苷酸替换模型通过 IQ-TREE v 1.0 (Lam *et al.*, 2015) 分析得到。利用 MrBayes v3.12 (Huelsenbeck and Ronquist, 2001) 进行贝叶斯分 析,共运行1千万代。每隔1000代取一次样,舍 去收敛之前的数据样本。利用 RAxML v8.2.9 (Stamatakis, 2006) 进行最大似然法分析,自展检 验 (bootstrap) 值设为1000。

2 结果与分析

2.1 大头隆胸长蝽线粒体基因组结构

大头隆胸长蝽的线粒体基因组是双链闭合环状 DNA 大分子, GenBank 序列号为 MN857166。测得的长度为 14 562 bp,由于控制区的特殊结构碱基组成,没有测得控制区的全部序列。已测得的序列包含一部分控制区和典型的 37 个基因,包括 22 个转运 RNA 基因,13 个蛋白编码基因和 2 个核糖体 RNA 基因(图1,表3)。大头隆胸长蝽线粒体基因组的基因组成和排列方式保守,与果蝇 Drosophila yakuba (双翅目:果蝇科)和大多数蝽次目昆虫相同(Clary and Wolstenholme,1985; Hua et al., 2008)。

大头隆胸长蝽线粒体基因组内发生基因重叠 的区域共有 15 处,总长度为 32 bp,大小从 1 ~ 7 bp 不等。最长 2 处基因重叠为 7 bp,分别位于 *ATP8*和*ATP6*,*ND4*和*ND4L*之间,并且这两段重 叠序列互为反向互补序列(ATGATAA),这与 Cameron(2014a)中报道的重叠序列相同。最短 的 10 处基因重叠为 1 bp,分别位于 *tRNA-Gln*和 *tRNA-Met、tRNA-Trp*和*tRNA-Cys、ATP6*和*COIII*、

总科 Superfamily	科 Family	种 Species	GenBank 序列号 Accession number
蝽总科 Pentatomoidea	兜蝽科 Dinidoridae	Megymenum gracilicorne	NC_042810
	龟蝽科 Plataspidae	Megacopta cribraria	NC_015342
红蝽总科 Pyrrhocoroidea	红蝽科 Pyrrhocoridae	Dysdercus cingulatus	NC_012421
		Dindymus rubiginosus	NC_042439
	大红蝽科 Largidae	Physopelta cincticollis	NC_042433
		Physopelta gutta	NC_012432
缘蝽总科 Coreoidea	姬缘蝽科 Rhopalidae	Stictopleurus subviridis	NC_012888
		Aeschyntelus notatus	NC_012446
	缘蝽科 Coreidae	Anoplocnemis curvipes	NC_035509
		Notopteryx soror	NC_037376
	蛛缘蝽科 Alydidae	Riptortus pedestris	NC_012462
长蝽总科 Lygaeoidea	跷蝽科 Berytidae	Yemmalysus parallelus	NC_012464
		Metatropis longirostris	NC_037373
	束蝽科 Colobathristidae	Phaenacantha marcida	NC_012460
	束长蝽科 Malcidae	Malcus inconspicuus	NC_012458
		Chauliops fallax	NC_020772
	大眼长蝽科 Geocoridae	Geocoris pallidipennis	NC_012424
	长蝽科 Lygaeidae	Kleidocerys resedae resedae	KJ584365
		Lygaeus sp. FS-2019	MF497725
	地长蝽科 Rhyparochromidae	Panaorus albomaculatus	NC_031364
		Eucosmetus incisus	MN857166

表 2 本研究中所使用的类群信息 Table 2 General information of the taxa used in this study

ND3 和 tRNA-Ala、tRNA-Asn 和 tRNA-Ser (GCU)、 tRNA-Ser (GCU) 和 tRNA-Glu、tRNA-Phe 和 ND5、 ND4 和 ND4L、ND6 和 CytB 以及 CytB 和 tRNA-Ser (UGA) 之间。同时观察到基因间隔区域有 6 处, 总长度为 34 bp,大小从 1~17 bp 不等,最长的基 因间隔发生在 tRNA-Ser (UGA) 和 ND1 之间,为 17 bp。37 个编码基因中,有9 个蛋白质编码基因 和 14 个 tRNA 基因由 J 链编码,4 个蛋白质编码基 因和 8 个 tRNA 基因以及 2 个 rRNA 基因由 N 链 编码。

2.2 蛋白质编码基因

大头隆胸长蝽的线粒体基因组包含 13 个蛋白 质编码基因,全长共 10 937 bp。除起始和终止子 之外,共发现 3 622 个密码子。其中 *ND1、ND4、 ND4L、ND5* 位于 N 链,而其余 9 个蛋白质编码基 因位于 J 链。A + T 含量为 77.9%,G+C 含量为 22.1%。除了 COI 使用 TTG 作为起始密码子以外, 其它所有蛋白质编码基因都以 ATN 作为起始密码 子。其中, ND2、 COII、 ND5、 ND4L、 ND6 和 NDI 使用 ATT 作为起始密码子, ATP6、 COII、 ND4 和 CytB 使用 ATG 作为起始密码子, ATP8 和 ND3 使用 ATA 作为起始密码子。这种非传统的 COI 起始密码子在蝽次目昆虫中非常常见。终止密 码子方面, 有 10 个蛋白质编码基因使用最常见的 三联体密码(TAA 和 TAG) 作为为终止密码子, 其中, ND5 使用 TAG 作为终止密码子,其余9个 使用 TAA 作为终止密码子。另外, NDI 使用 TA 作为终止密码子, COII 和 COIII 的则使用单个T 作 为终止密码子(表4)。以 TA 或T 作为终止密码 子的现象在昆虫线粒体基因组中很常见,有研究 推测完全终止密码子 TAA 可以通过转录后多腺苷 酸化产生(Ojala *et al.*, 1981)。



图1 大头隆胸长蝽线粒体基因组结构

Fig. 1 Mitochondrial genome map of Eucosmetus incisus

注: 箭头表示基因转录的方向。蓝色代表蛋白质编码基因,红色代表 tRNA 基因,紫色代表 rRNA 基因,灰色代表已测 得控制区,斜线代表未测得的控制区。Note: Arrows indicated the orientation of gene transcription. PCGs were showed as blue arrows, tRNA genes as red arrows, rRNA genes as purple arrows, the sequenced control region as gray arrows and the unsequenced control region as slash.

2.3 tRNA

大头隆胸长蝽线粒体基因组包含典型的 22 个 红蝽科) (Men *e* tRNA 基因,其中有 14 个位于 J 链上,8 个位于 N 大头隆胸长 链上,长度范围在 62 bp (tRNA-Asp 和 tRNA-Gly) 码子环具有极低 到 73 bp (tRNA-Lys) 之间。其 tRNA 二级结构比较 子臂长度比较保 保守,除 tRNA-His 之外的所有 tRNA 都能够折叠成 4 bp 外,其余 tB 经典的三叶草的二级结构(图 2)。根据 Mitos DHU 臂长度为 2 WebServer 的分析,在 tRNA-His 的二级结构中 TψC 化量最大的是 DI 臂的 "环"结构缺失,仅有 "茎"结构,这一现象 1~8 bp (表 5)。

在巨红蝽 Macrocheraia grandis grandis (半翅目: 大 红蝽科) (Men et al., 2019) 中同样有报道。

大头隆胸长蝽的 tRNA 的氨基酸接受臂和反密 码子环具有极低的变异性,长度都为 7 bp。反密码 子臂长度比较保守,除了 tRNA-Arg 和 tRNA-Glu 为 4 bp 外,其余 tRNA 的反密码子臂长度都为 5 bp。 DHU 臂长度为 2 ~ 4 bp, T ψ C 臂长度为 3 ~ 5 bp。变 化量最大的是 DHU 环和 T ψ C 环,分别为 4 ~ 8 bp 和 1 ~ 8 bp (表 5)。

基因 Gene	链 Strand	位置 Position	长度 (bp) Length	反密码子 Anticodon	起始密码子 Start codon	终止密码子 Stop codon	间隔 Intergenic
		1 ((C A T		1	nucleotides
tRNA-le	+	1 - 66	66	GAI			2
tRNA-Gln	-	64 - 132	69	TIG			- 3
tRNA-Met	+	132 – 199	68	CAT			- 1
ND2	+	200 – 1168	969		АТТ	ТАА	0
tRNA-Trp	+	1167 – 1229	63	ТСА			-2
tRNA-Cys	-	1229 – 1293	65	GCA			- 1
tRNA-Tyr	-	1294 – 1356	63	GTA			0
COI	+	1358 – 2896	1 534		TTG	TAA	1
tRNA-Leu (UAA)	+	2892 - 2955	64	TAA			-5
COII	+	2956 - 3622	667		ATT	Τ –	0
tRNA-Lys	+	3623 - 3695	73	CTT			0
tRNA-Asp	+	3696 - 3757	62	GTC			0
ATPase8	+	3758 - 3913	156		ATA	TAA	0
ATPase6	+	3907 - 4569	663		ATG	TAA	-7
COIII	+	4569 - 5355	787		ATG	Т –	- 1
tRNA-Gly	+	5356 - 5417	62	TCC			0
ND3	+	5418 - 5768	351		ATA	TAA	0
tRNA-Ala	+	5768 - 5830	63	TGC			- 1
tRNA-Arg	+	5832 - 5894	63	TCG			1
tRNA-Asn	+	5896 - 5961	66	GTT			1
tRNA-Ser (GCU)	+	5961 - 6029	69	GCT			- 1
tRNA-Glu	+	6029 - 6092	64	TTC			- 1
tRNA-Phe	-	6093 - 6156	64	GAA			0
ND5	-	6156 - 7853	1 698		ATT	TAG	- 1
tRNA-His	-	7854 – 7919	66	GTG			0
ND4	-	7919 – 9235	1 317		ATG	TAA	- 1
ND4L	-	9229 - 9507	279		ATT	TAA	-7
tRNA-Thr	+	9510 - 9574	65	TGT			2
tRNA-Pro	-	9575 - 9637	63	TGG			0
ND6	+	9640 - 10101	462		ATT	TAA	2
CytB	+	10101 – 11234	1 134		ATG	TAA	- 1
tRNA-Ser (UGA)	+	11234 - 11302	69	TGA			- 1
ND1	-	11320 - 12239	920		ATT	TA –	17
tRNA-Leu (UAG)	-	12240 - 12304	65	TAG			0
16S rRNA	-	12305 - 13561	1 257				0
tRNA-Val	_	13562 - 13628	67	TAC			0
12S rRNA	-	13629 - 14435	807				0
Control region		14436 - 14562	127				

表 3 大头隆胸长蝽线粒体基因组结构

 Table 3 Organization of Eucosmetus incisus mitochondrial genome

Table 4 Start and stop codons of 13PCGs in Eucosmetus incisus mitochondrial genome 密码子 Codon 蛋白质编码基因 PCGs 密码子 Codon 蛋白质编码基因 PCGs 起始密码子 Start codon 终止密码子 Stop codon ATT ND2 , COII , ND5 , ND4L , ND6 , ND1 ND2, COI, ATP8, ATP6, ND3, ND4, ND4L, ND6, CytB TAA ATA ATP8 , ND3 TAG ND5 ATG ATP6 , COII , ND4 , CytB TA COII, COIII TTG Т COI ND1 J链 J strand Ala (A) A ^A AAU 111 A .. UUA uco GUU AUC III GUUA UUA ucc CUA υ^υ 6-6-0central de la constante de la IIGA ugu ec.u N链 N strand UA^{UGAI} N GIIG 氨基酸接受臂 acid acc DHU臂 DHU st UUA UUU AA/ DHU环 DHUI 反密码子臂

大头隆胸长蝽线粒体基因组中 tRNA 结构 图 2

Fig. 2 Predicted secondary structures of tRNAs in Eucosmetus incisus mitochondrial genome

反密码子环

N

注: tRNAs 使用相应的氨基酸缩写表示。Watson-Crick 经典配对用短线表示,GU 配对用星号表示,其他非经典配对用空 心圆表示。Note: The tRNAs were labeled with the abbreviations of their corresponding amino acids. Inferred Watson-Crick bonds were illustrated by lines , GU bonds by asterisk and the other non-Watson-Crick interactions were represented by hollow circles.

บ ัฐบุล

ប៉ិ បឲឲ

表4 大头隆胸长蝽线粒体基因组中蛋白质编码基因的起始密码子和终止密码子

	14010			/// ··· ···					
tRNA	链 Strand	氨基酸 接受臂 (bp) AA stem	反密码 子臂 (bp) AC stem	反密码 子环 (nucleotide) AC loop	DHU 臂 (bp) DHU stem	DHU 环 (nucleotide) DHU loop	TΨC 臂 (bp) TΨC stem	ТΨС 环 (nucleotide) ТΨС loop	可变环 (nucleotide) Variable loop
tRNA-Ala	J	7	5	7	4	5	3	5	4
tRNA-Arg	J	7	4	7	4	4	5	1	5
tRNA-Asn	J	7	5	7	3	7	3	7	5
tRNA-Asp	J	7	5	7	4	5	3	4	4
tRNA-Glu	J	7	4	7	4	6	4	3	5
tRNA-Gly	J	7	5	7	2	7	3	4	4
tRNA-Ile	J	7	5	7	3	6	4	6	5
tRNA-Leu (_UAA)	J	7	5	7	3	6	4	5	4
(UAA) tRNA-Lys	J	7	5	7	4	8	4	9	5
tRNA-Met	J	7	5	7	4	5	5	6	4
tRNA–Ser	J	7	5	7	3	4	5	9	4
RNA-Ser	J	7	5	7	3	6	5	7	4
(UGA) tRNA–Thr	J	7	5	7	4	5	3	7	4
tRNA-Trp	J	7	5	7	4	5	4	3	4
tRNA-Cys	Ν	7	5	7	4	5	4	5	4
tRNA-Gln	Ν	7	5	7	3	5	5	7	4
tRNA-His	Ν	7	5	7	4	7	5	2	4
tRNA-Leu	Ν	7	5	7	3	6	5	4	4
(UAG) tRNA-Phe	Ν	7	5	7	4	7	3	4	4
tRNA-Pro	Ν	7	5	7	4	5	3	5	4
tRNA-Tyr	Ν	7	5	7	3	6	4	4	4
tRNA-Val	Ν	7	5	7	2	4	4	8	5

表 5 大头隆胸长蝽线粒体基因组中 22 个 tRNA 的核苷酸分布情况 Table 5 Nucleotides distribution in 22 tRNAs of *Eucosmetus incisus* mitochondrial genome

此外,共发现 17 处非 Watson-Crick 碱基配对 存在于大头隆胸长蝽线粒体 tRNA 基因二级结构 中,且都为G=U配对,其中有15个集中在氨基 酸接受臂和DHC臂上,剩余2个分别位于反密码 子臂和 T ψ C臂上(表6)。

2.4 rRNA

大头隆胸长蝽的线粒体基因组中的 16S rRNA 基因长 1 257 bp,位于 *tRNA-Leu*(*UAG*)和 *tRNA-Val*之间,其二级结构包含 6 个结构域(节肢动物 中结构域 III 缺失)和 45 个茎环结构(图 3)。和 长蝽总科内的其它物种相比,大头隆胸长蝽在 H991、H1196、H235、H2735 以及 H183 到 *tRNA-* Val 之间的茎环结构存在较大变异。而 H1775、 H2064、H2507 等二级结构则和长蝽总科其它物种 相比,无论在序列还是在二级结构上都十分保守。 这与已报道的长蝽总科线粒体基因组内 16S rRNA 的结构域 IV 和 V 比结构域 I、II、VI 更保守的研 究结果一致(Li *et al.*,2013; 2016ab)。

12S rRNA 基因长 807 bp, 位于 *tRNA-Val* 和控制区之间,其二级结构包括 3 个结构域和 27 个茎环结构(图 4),发现相比结构域 I 和 II,结构域III 更保守。特别是在结构域 II 中,从 H567 到H769存在一个拉链状二级结构,大头隆胸长蝽的这部分结构与长蝽总科其它物种相比,不仅在序

		-		0	•
tRNA	链 Strand	氨基酸接受臂 (bp) AA stem	反密码子臂(bp) AC stem	DHU 臂 (bp) DHU stem	TΨC 臂 (bp) TΨC stem
tRNA-Ala	J	(G – U) (G – U)			
tRNA-Asp	J			(G – U)	
tRNA–Ser (GCU)	J				(G – U)
tRNA-Cys	Ν	(G – U) (G – U)		(G – U)	
tRNA-Gln	Ν	(G – U)			
tRNA-His	Ν	(G – U) (G – U)			
tRNA-Leu (UAG)	Ν		(G – U)		
tRNA-Phe	Ν	(G – U)		(G – U)	
tRNA-Pro	Ν	(G – U)		(G – U)	
tRNA-Tyr	Ν			(G – U)	
tRNA-Val	Ν			(G – U)	







139



图 4 大头隆胸长蝽线粒体基因组中 12S rRNA 二级结构

Fig. 4 Predicted secondary structure of the 12S rRNA in Eucosmetus incisus mitochondrial genome

列长度上存在差异,而且在核苷酸组成上也存在 较高的碱基替换,使得这部分茎和环的长度各有 不同,但整体上仍保持一个拉链状结构。rRNA的 同源性更多的体现在二级结构保守性上,而不是具 有某一段保守的序列,说明rRNA结构构成上的生 物学意义要大于它的序列组成(陈国忠等,2005)。 2.5 核苷酸组成和密码子使用

在所测得的大头隆胸长蝽线粒体基因组中, ATCG 4 种核苷酸含量分别为 45.9%、32.4%、 13% 和 8.6%。AT 偏斜率(AT-skew)大小为 17%,GC 偏斜率(GC-skew)为-20%,具有明显 的 AT 偏向性。这与大部分其它蝽次目昆虫类似, AT-skew 基本为正值,即整体上A 的含量大于T, GC-skew 基本为负值,表明 C 的含量大于 G。大头 隆胸长蝽线粒体基因组总体上具有较高的 A + T 含 量,为 78.3%,在蛋白质编码基因中为 77.9%, 在 tRNA 基因中为 79.1%,rRNA 基因的 A + T 含 量最高,为 80%。在蛋白质编码基因中,A + T 含 量最高的基因是 *ND*6(87.8%),含量最低的基因 是 *COI*(72.1%)。总基因组、蛋白质编码基因 J 链表现为 AT 偏移和 CG 偏移,总蛋白质编码基因 J 链表现为 AT 偏移和 CG 偏移,总蛋白质编码基因 X 链则刚好相反,表现为 TA 偏移和 GC 偏移,总 tRNA 编码基因及其 J 链表现为 AT 偏移和 GC 偏移 (表 7)。由此可见,线粒体基因组中核苷酸组成在 不同链间是不对称的。 表7 大头隆胸长蝽线粒体基因组核苷酸组成分析

Table 7 Nucleo	otide composit	ion of the	e Eucosmet	tus incisus	mitochond	lrial genom	e	
结构 Feature	长度 (bp) Length	T 含量 T%	C 含量 C%	A 含量 A%	G 含量 G%	AT 含量 A + T%	AT 偏移 AT-Skew	GC 偏移 GC-Skew
全线粒体基因组 Whole genome	14 562	32.4	13.0	45.9	8.6	78.3	0. 17	-0.20
蛋白质编码基因 Protein coding genes	10 937	43.5	10.6	34.4	11.5	77.9	-0.12	0.04
J 链蛋白质编码基因 Protein coding genes-J	6 723	37. 1	12.7	40. 1	10. 1	77.2	0.04	-0.12
N 链蛋白质编码基因 Protein coding genes-N	4 214	53.7	7.2	25.4	13.7	79. 1	- 0. 36	0. 31
蛋白质编码基因第一位密码子 First codon position	3 646	38.4	10. 9	35.8	14. 8	74.2	- 0. 04	0. 15
蛋白质编码基因第二位密码子 Second codon position	3 646	46.0	14. 6	27.0	12.4	73.0	- 0. 26	- 0. 08
蛋白质编码基因第三位密码子 Third codon position	3 645	46. 1	6.3	40.4	7.2	86.5	- 0. 07	0.06
转运 RNS 基因 tRNA genes	1 439	39.0	8.9	40.1	12.0	79.1	0.01	0.15
J 链转运 RNS 基因 tRNA genes-J	917	36.5	10.5	42.1	10.9	78.6	0.07	0.02
N 链转运 RNS 基因 tRNA genes-N	522	43.3	6.1	36.6	14.0	79.9	- 0. 08	0.39
核糖体 RNA 基因 rRNA genes	2 064	50.9	6.6	29.1	13.4	80.0	-0.27	0.34
ATP6	663	36.5	12.8	42.2	8.4	78.7	0.07	-0.21
ATP8	156	38.5	7.7	49.4	4.5	87.8	0.12	-0.26
COI	1 539	37.2	14.0	34.9	13.9	72. 1	-0.03	0.00
COII	667	35.8	13.3	40.8	10.0	76.6	0.06	-0.14
COIII	787	36.6	13.3	37.4	12.7	74.0	0.01	-0.02
CytB	1 134	35.0	15.0	39.6	10.4	74.6	0.06	-0.18
ND1	920	53.9	8.0	23.4	14.7	77.3	-0.40	0. 29
ND2	969	41.8	9.2	42.1	6.9	83.9	0.00	-0.14
ND3	351	37.9	12.3	42.7	7.1	80.6	0.06	-0.26
ND4	1 317	53.1	7.5	25.3	14. 1	78.4	-0.35	0.31
ND4L	279	59.5	4.7	22. 2	13.6	81.7	-0.46	0. 49
ND5	1 698	53.2	6.9	27.0	12.8	80.3	-0.33	0.30
ND6	462	34.8	10.6	49.6	5.0	84.4	0.17	-0.36
12S rRNA	807	49.4	6.9	29.1	14.5	78.6	-0.26	0.35
16S rRNA	1 257	51.9	6.4	29.0	12.6	80.9	-0.28	0.33

值得注意的是,由N链编码的蛋白质基因和 rRNA基因都是TA偏移和GC偏移,而除COI之 外所有由J链编码的蛋白质基因都刚好相反,都是 AT 偏移和 CG 偏移, COI为 TA 偏移和 CG 偏移 (图 5)。昆虫线粒体基因组 J 链、N 链上蛋白质编 码基因核苷酸组成偏向性刚好相反的情况在 C. *fallax*(Li *et al.*, 2013) 也有发现,有报道称 GC-skew 和 AT-skew 的值因复制起点的方向和密码子 位置的变化而改变(Hassanin *et al.*, 2005; Wei *et al.*, 2010)。为了深入了解这种现象的机制,需 要进行更多关于线粒体基因组序列和功能的研究 工作。

核苷酸的 AT 偏向性也反映在密码子使用中, 蛋白质编码基因密码子的使用表现出极大的不均 质性,其中密码第三位的 A + T 含量最高,为 86.5%,第一位和第二位次之,分别为 74.2% 和 73%。密码子 ATT (371)、TTA (356)、ATA (330) 以及 TTT (309) 是使用频率最高的四种密 码子,全部由 A、T 构成,分别转运氨基酸 Ile、 Leu、Met 和 Phe。对于大多数氨基酸来说,使用 最频繁的密码子是 NNA 和 NNU,而不是与 tRNA 反密码子严格配对的密码子(图6)。如甲硫氨酸 Met 对应的密码子是 AUA 和 AUG,使用次数分别 为 330 和 29,其中 AUG 是与 *tRNA-Met* 的反密码子 严格配对的密码子。



图 5 大头隆胸长蝽线粒体基因组的 AT 偏移和 GC 偏移



2.6 系统发育

本研究选取4个总科共21个物种,根据含有 蛋白质编码基因全部123位密码子的矩阵 PCG123, 和去除第三位密码子的矩阵 PCG12,分别进行最 大似然法和贝叶斯分析,得到了不同的系统发育 结果。其中,通过矩阵 PCG123 得到的两个系统发 育树完全不相同,并且与客观事实不尽相符,而 通过矩阵 PCG12 分析得到具有相同拓扑结构的系 统发育树。通过对比有无密码子第3位碱基得到 的系统发育结果的拓扑结构是否一致,发现第3 位密码子的碱基组成异质性较高,容易形成核苷 酸替代饱和,严重影响树形结构,表明密码子第3 位的存在与否对于系统发育研究非常重要。

根据矩阵 PCG12 得到系统发育树: { [(长蝽 总科+缘蝽总科)+红蝽总科]+蝽总科},这个结 果很好地支持了长蝽总科的单系性,在总科关系 方面,长蝽总科和缘蝽总科亲缘关系更近 (图7)。这与 Xie *et al.* (2005)和 Hua *et al.* (2008)根据 18S rRNA 和线粒体基因组数据得到的树形一致。

从贝叶斯分析来看,关于长蝽总科、缘蝽总 科、红蝽总科、蝽总科系统发育分支节点处的后 验概率均为100%,而在长蝽总科内部,(地长蝽 科+长蝽科)与束长蝽的分支节点处后验概率为 85%,小于95%这一通常认为的较为可靠的阈值 (Leaché and Reeder,2002; Suzuki *et al.*,2002)。 从自展检验的值来看,4个总科的分支节点支持率 均高于 70% 这一通常认为的可靠的阈值 (Huelsenbecket *et al.*,1993),长蝽总科内部部分 节点支持率偏低。



图 6 大头隆胸长蝽线粒体基因组中每个氨基酸的同义密码子使用率

Fig. 6 Percentage of synonymous codon usage of each amino acid in the *Eucosmetus incisus* mitochondrial genome 注: X 轴上为密码子家族。Note: Codon families are provided on the x-axis.

这说明各总科间的系统发育关系总的来讲得 到了解析,但还需要进一步去验证;而长蝽总科 内部关系仍处于未解析的状态。造成这种现象的 原因可能有两个,一是长蝽总科内部取样不充分, 二是仅根据线粒体基因组进行系统发育推断对于 类群选取比较敏感。

3 结论与讨论

本研究测定了大头隆胸长蝽编码区域的基因 序列,并分析了该物种线粒体基因组特征,这是 对缢胸族昆虫线粒体基因组序列的首次报道。大 头隆 胸长 蝽线粒体基因组已测得部分大小为 14 562 bp,包含标准的13个蛋白编码基因,22个 转运 RNA 基因,2个核糖体 RNA 基因以及部分控 制区。其线粒体基因排列顺序同亚库巴果蝇 Drosophila yakuba 和大多数蝽次目昆虫排列顺序相 同。线粒体基因组序列存在基因重叠和基因间隔 现象,并发现有两个长度为7 bp 的基因重叠片段 互为反向互补序列(ATGATAA),一个位于 ATP8 和 ATP6 之间,另一个位于 ND4 和 ND4L 之间。蝽 次目昆虫中大部分物种在 tRNA-Ser(UGA) 和 ND1 之间存在长度不等的非编码区。

线粒体基因组中核苷酸组成在不同链间不对称,大头隆胸长蝽线粒体核苷酸组成表现出了很高的 AT 偏向性,13 个蛋白编码基因和 2 个核糖体 RNA 中,由 N 链编码的基因都是 TA 偏移和 GC 偏移,而除 COI 之外所有由 J 链编码的基因都刚好相反,都是 AT 偏移和 CG 偏移,COI 为 TA 偏移和 CG 偏移,该观察结果很可能归因于复制方向的不 对称。此外,核苷酸 AT 偏向性也反映在密码子的使用中,使用最频繁的密码子均由 AT 组成,且多数并不与 tRNA 反密码子严格配对。

非传统起始密码子 TTG 在蝽次目昆虫中广泛





Fig. 7 Phylogenetic tree inferred from the sequences of PCGs in mitochondrial genome 注: 贝叶斯和最大似然法分析具有一致的拓扑结构。节点处上方的数字代表贝叶斯后验概率,下方代表 bootstrap

fa。Note: Bayesian analyses and Maximum likelihood showed the same topology. Numbers at the nodes were Bayesian posterior probabilities (up) and Maximum likelihood bootstrap values (down) $\,$.

存在,通常发生在蛋白质编码基因 COI 中,以单 个T 作为终止密码子的现象多发生在 COII 和 COIII 中,而以 TA 作为终止密码子的情况仅随机出现在 少部分基因中。大头隆胸长蝽线粒体基因中,除 *tRNA-His* 因缺少 TψC 环,不能正常折叠外,其它 21 个 tRNA 均能折叠成经典三叶草结构。大头隆 胸长蝽的 rRNA 二级结构在一些非保守结构域与长 蝽总科其它物种相比,存在长度差异和较高的碱 基替换,但并不影响各物种具有大致相同的二级 结构,说明 rRNA 结构构成上的生物学意义要大于 它的序列组成。

长期以来,长蝽总科、红蝽总科、缘蝽总科 三者间的系统发育关系一直存在争议,且至今尚 无较为统一的意见。本研究的系统发育结果是基 于蛋白质编码基因的第1、2位密码子构成的数据 集得到的,系统发育结果很好地确认了长蝽总科 的单系性,并且支持(((长蝽总科+缘蝽总科) +红蝽总科) +蝽总科)这一关系。而长蝽总科 内部关系仍处于未解析状态,可能原因是长蝽总 科内部取样不充分,另外仅根据线粒体基因组进 行系统发育推断对于类群选取比较敏感,因此未 来需要尽可能选取完整的科级分类单元类群,并 结合更多分子标记(核基因分子标记)和形态学 来较为全面的进行探讨和研究。大头隆胸长蝽线 粒体基因组的测序不仅为确定长蝽总科、缘蝽总 科和红蝽总科的系统发育关系提供更多证据,还 为将来开展地长蝽科缢胸族昆虫相关的分子系统 发育研究初步提供了基础数据。

参考文献 (References)

- AminN, Daha L, Agus N, et al. Diversity of some endophytic fungi associated with rice black bug Paraeucosmetus pallicornis on rice plant [J]. Journal of Chemical and Pharmaceutical Research, 2015, 7 (4): 1246-1253.
- Bernt M , Donath A , Jühling F , et al. MITOS: Improved de novo metazoan mitochondrial genome annotation [J]. Molecular Phylogenetics and Evolution , 2013 , 69 (2): 313 – 319.
- Boore JL. Complete mitochondrial genome sequence of the polychaete annelid Platynereis dumerilii [J]. Molecular Biology and Evolution, 2001, 18 (7): 1413-1416.
- Boore JL. Animal mitochondrial genomes [J]. Nucleic Acids Research, 1999, 27 (8): 1767-1780.
- Cameron SL, Whiting MF. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta* (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability

within butterflies and moths [J]. Gene , 2008 , 408 (1) : 112 – 123.

- Cameron SL. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research [J]. Systematic Entomology , 2014a , 39 (3): 400-411.
- Cameron SL. Insect mitochondrial genomics: Implications for evolution and phylogeny [J]. Annual Review of Entomology, 2014b, 59: 95-117.
- Cannone JJ, Subramanian S, Schnare MN, et al. The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs [J]. BMC Bioinformatics, 2002, 3 (1): 2.
- Chen GZ, Li WJ, Xu LH, et al. Advanced in 16S rRNA secondary structure and its application in systematic classification [J]. Journal of Microbiology, 2005, 25 (5): 54 – 57. [陈国忠,李文 均,徐丽华,等. 16S rRNA 二级结构的研究进展及其在系统分 类中的应用 [J]. 微生物学杂志, 2005, 25 (5): 54 – 57]
- Chen M , Tian LL , Shi QH , et al. Complete mitogenome of the Lesser Purple Emperor Apatura ilia (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies [J]. Zoological Research , 2012 , 33: 191 – 201.
- Clary DO, Wolstenholme DR. The mitochondrial DNA molecule of Drosophila yakuba: Nucleotide sequence, gene organization, and genetic code [J]. Journal of Molecular Evolution, 1985, 22 (3): 252 – 271.
- Gillespie JJ, Johnston JS, Cannone JJ, et al. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of Apis mellifera (Insecta: Hymenoptera), structure, organization and retrotransposable elements [J]. Insect Molecular Biology, 2006, 15 (5): 657-686.
- Gissi C , Iannelli F , Pesole G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species [J]. *Heredity* , 2008 , 101: 301 – 320.
- Grant JR , Stothard P. The CG view server: A comparative genomics tool for circular genomes [J]. Nucleic Acids Research , 2008 , 36: 181 – 184.
- Hassanin A , Leger N , Deutsch J. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa , and consequences for phylogenetic inferences [J]. Systematic Biology , 2005 , 54 (2): 277 – 298.
- Henry TJ. Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea [J]. Annals of the Entomological Society of America, 1997, 90 (3): 275 – 301.
- Hua JM, Li M, Dong PZ, et al. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera) [J]. BMC Genomics, 2008, 9: 610.
- Huelsenbeck JP , Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees [J]. *Bioinformatics* , 2001 , 17: 754 – 755.
- Huelsenbeck JP , Hillis DM. Success of phylogenetic methods in the fourtaxon case [J]. Systematic Biology , 1993 , 42 (3) : 247 – 264.

- Hwang UW, Park CJ, Yong TS, et al. One step PCR amplification of complete arthropod mitochondrial genomes [J]. Molecular Phylogenetics and Evolution, 2001, 19 (3): 345 – 352.
- KumarS, 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.
- Lalitha S. Primer Premier 5 [J]. Biotech Software and Internet Report, 2000, 1 (6): 270 – 272.
- Lam TN, Schmidt HA, Arndt VH, et al. IQ TREE: A fast and effective stochastic algorithm for estimating maximum – likelihood phylogenies [J]. Molecular Biology and Evolution, 2015, 32 (1): 268 – 274.
- Leaché AD, Reeder TW. Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): A comparison of parsimony, likelihood, and Bayesian approaches [J]. *Systematic Biology*, 2002, 51 (1): 44 – 68.
- Li T, Gao CQ, Cui Y, et al. The complete mitochondrial genome of the stalk – eyed bug *Chauliops fallax* Scott, and the monophyly of Malcidae (Hemiptera: Heteroptera) [J]. *PLoS ONE*, 2013, 8 (2): e55381.
- Li T, Yang J, Li YW, et al. A mitochondrial genome of Rhyparochromidae (Hemiptera: Heteroptera) and a comparative analysis of related mitochondrial genomes [J]. Scientific Reports, 2016a, 6: 35175.
- Li T, Yi WB, Zhang HG, et al. Complete mitochondrial genome of the birch catkin bug Kleidocerys resedae resedae, as the first representative from the family Lygaeidae (Hemiptera: Heteroptera: Lygaeoidea) [J]. Mitochondrial DNA, 2016b, 27 (1): 618 – 619.
- Lowe TM , Chan PP. tRNA scan SE On line: Integrating search and context for analysis of transfer RNA genes [J]. Nucleic Acids Research , 2016 , 44: 54 – 57.
- Ma C , Liu C , Yang P , et al. The complete mitochondrial genomes of two band – winged grasshoppers , Gastrimargus marmoratus and Oedaleus asiaticus [J]. BMC Genomics , 2009 , 10: 156.
- Mathews DH, Reuter JS. RNA structure: Software for RNA secondary structure prediction and analysis [J]. BMC Bioinformatics, 2010, 11 (1): 873.
- Men Y, Kment P, Stahlavsky F, et al. The mitochondrial genomes of Macrocheraia grandis grandis and Myrmoplasta mira and the unique mitogenome rearrangement in Pyrrhocoroidea (Hemiptera: Heteroptera: Pentatomomorpha) [J]. Entomotaxonomia, 2019, 41 (2): 96 – 113.
- Ojala D , Montoya J , Attardi G. tRNA punctuation model of RNA processing in human mitochondria [J]. *Nature* , 1981 , 290: 470 474.
- Perna NT , Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes [J]. Journal of Molecular Evolution , 1995 , 41 (3): 353 – 358.
- Reineke A , Karlovsky P , Zebitz CP. Preparation and purification of DNA from insects for AFLP analysis [J]. Insect Molecular Biology , 2010 , 7 (1): 95 – 99.

- Rosmana A, Sjam S, Sari DE, et al. Fungi associated with Paraeucosmetus pallicornis causing apparent symptoms of toxicity in rice grain and rice seedlings [J]. International Journal of Current Microbiology and Applied Sciences, 2014, 3 (2): 407-414.
- Schaefer CW. Heteropteran trichobothria (Hemiptera: Heteroptera) [J]. International Journal of Insect Morphology and Embryology, 1975, 4: 193-264.
- Schuh RT and Slater JA. True Bugs of the World (Hemiptera: Heteroptera), Classification and Natural History [M]. New York: Cornell University Press, 1995, 229 – 233.
- Song HY, Yu YB, Dong JY, et al. Catalogue of rice insect pests in Hunan Province [J]. Journal of Hunan Agricultural College, 1986,
 4: 35-49. [宋慧英,余映波,董浚玙,等. 湖南省水稻害虫名
 录 [J]. 湖南农学院学报, 1986, 4: 35-49]
- Stamatakis A. RAxML VI HPC: Maximum likelihood based phylogenetic analyses with thousands of taxa and mixed models [J]. *Bioinformatics*, 2006, 22 (21): 2688 - 2690.
- Suzuki Y , Glazko GV , Nei M. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics [J]. Proceedings of the National Academy of Sciences , 2002 , 99 (25): 16138 – 16143.
- Swindell SR , Plasterer TN. SEQMAN. Contig assembly [J]. Methods in Molecular Biology , 1997 , 70: 75 – 89.

Taanman JW. The mitochondrial genome: Structure, transcription,

translation and replication $[\,J\,].$ Biochimica et Biophysica Acta , 1999 , 1410: 103 – 123.

- Timmermans MJ, Dodsworth S, Culverwell CL, et al. Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics [J]. Nucleic Acids Research, 2010, 38: E197.
- Tullgren A. Zur morphologie und systematik der Hemipteran [J]. Entomologisk Tidskrift, 1918, 39: 113 – 132.
- Wei SJ, Shi M, Chen XX, et al. New views on strand asymmetry in insect mitochondrial genomes [J]. PLoS ONE, 2010, 5: E12708.
- Xie Q, Bu WJ, Zheng LY. The Bayesian phylogenetic analysis of the 18S rRNA sequences from the main lineages of Trichophora (Insecta: Heteroptera: Pentatomomorpha) [J]. Molecular Phylogenetics and Evolution, 2005, 34 (2): 448-451.
- Yuan ML, Zhang QL, Guo ZL, et al. Comparative mitogenomic analysis of the superfamily Pentatomoidea (Insecta: Hemiptera: Heteroptera) and phylogenetic implications [J]. BMC Genomics, 2015, 16: 460.
- Zheng LY, Zou HG. Chinese Bug Identification Manual (Hemiptera: Heteroptera) [M]. Beijing: Science Press, 1981: 161 – 184.
 [郑乐怡,邹环光.中国蝽类昆虫鉴定手册(半翅目:异翅亚目) [M].北京:科学出版社, 1981: 161 – 184]