



麻竖毛天牛线粒体基因组测定及 天牛科线粒体基因组比较分析*

杨 锦^{1**} 曹利军² 耿涌鑫¹ 魏丹峰¹ 陈 敏^{1***}

(1. 北京林业大学省部共建森林培育与保护教育部重点实验室, 北京 100083;

2. 北京市农林科学院植物保护环境保护研究所, 北京 100097)

摘 要 【目的】线粒体基因组分析是研究昆虫系统发育的重要手段。本研究通过测定麻竖毛天牛 *Thyestilla gebleri* (Faldermann, 1835) 线粒体基因组, 比较分析天牛科 Cerambycidae 线粒体基因组的特征, 进而初步探讨麻竖毛天牛系统发育地位和天牛科部分类群之间的系统进化关系。【方法】采用引物步移法测定麻竖毛天牛线粒体基因组全序列。参照 GenBank 收录的 16 种天牛线粒体基因组序列进行基因注释; 采用在线软件 tRNAscan-SE Search Server 对转运 RNA (tRNA) 的二级结构进行了预测。通过对 16 种天牛的线粒体基因组进行重新注释, 结合本研究获得的麻竖毛天牛线粒体基因组进行序列特征比较分析。基于 11 个蛋白编码基因的核苷酸序列, 利用最大似然法构建了天牛科 17 种天牛的系统发育树。

【结果】麻竖毛天牛线粒体基因组全长 15 505 bp, A+T 含量为 74.07%, 包含 13 个蛋白编码基因 (PCGs), 2 个核糖体 RNA (rRNA) 基因, 22 个 tRNA 基因和一个长度为 872 bp 的控制区, 未发现基因重排。通过比对 17 种天牛的线粒体基因组, 发现长翅暗天牛 *Vesperus conicollis* 一个 tRNA (*trn P*) 基因的移位。17 种天牛的 tRNA 中, *trnS1* (AGN) 的 D 臂均缺失, 其余 tRNA 都具有典型的三叶草结构。大部分种类的蛋白编码基因起始密码子为典型的 ATN (ATA、ATT、ATC、ATG), 只有部分种类的 *Nad1*、*COI*、*ATP8* 基因存在特殊的起始密码子 (TTG、AAC、AAT、GTG), 终止密码子均为常见的 TAR (TAA、TAG) 或不完整的 T 和 TA。系统发育树中, 6 个亚科分别单独分支, 其中, 麻竖毛天牛与云斑白条天牛 *Batocera lineolata* 聚为一支。【结论】麻竖毛天牛线粒体基因组符合天牛科线粒体基因组的一般特征; 除少数 tRNA 基因存在重排外, 天牛科线粒体基因排列相对稳定; 基于线粒体基因组的系统发育分析支持天牛科 6 个亚科的单系性, 麻竖毛天牛和云斑白条天牛亲缘关系较近。

关键词 麻竖毛天牛, 天牛科, 线粒体基因组, 系统发育

Determination of the complete mitochondrial genome of *Thyestilla gebleri* and comparative analysis of the mitochondrial genome in Cerambycidae

YANG Jin^{1**} CAO Li-Jun² GENG Yong-Xin¹ WEI Dan-Feng¹ CHEN Min^{1***}

(1. The Key Laboratory of Silviculture and Forest Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China; 2. Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China)

Abstract [Objectives] Analysis of mitochondrial genome is an important way to reveal the phylogenetic relationship in insects. The aim of this study was to determine the mitochondrial genome of *Thyestilla gebleri* and comparatively analyze the characteristics of mitochondrial genome in Cerambycidae. We also investigated the phylogenetic position of *T. gebleri* and the phylogenetic relationship among species in Cerambycidae based on mitochondrial genomes. [Methods] The mitochondrial

*资助项目 Supported projects: 转基因生物新品种培育重大专项 (No. 2018ZX08021001)

**第一作者 First author, E-mail: 411823293@qq.com

***通讯作者 Corresponding author, E-mail: minch@bjfu.edu.cn

收稿日期 Received: 2017-01-22, 接受日期 Accepted: 2017-09-22

genome of *T. gebleri* was determined using primer walking method. The annotation of mitochondrial genes of *T. gebleri* were based on the GenBank data of mitochondrial genome in Cerambycidae. The secondary structure of transport RNA (tRNA) was predicted by online tRNAscan-SE Search Server. After re-annotating the mitochondrial genome of 16 species available in GenBank, we conducted comparative analysis of the mitochondrial genomic features for Cerambycidae. A phylogenetic tree for 17 species in Cerambycidae was constructed by Maximum Likelihood method based on the nucleotide sequences of 11 protein-coding genes. **[Results]** The total length of mitochondrial genome of *T. gebleri* was 15 505 bp with an A+T content of 74.07%. It contained 13 protein coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, 22 tRNA genes and A 872 bp-long control region, and no gene rearrangement was found. The results revealed the translocation of *trn P* (one of a tRNA gene) in *Vesperus conicicollis*. The secondary structure of tRNA genes had the typical clover structure except the absent of “D” arm in *trnS1*(AGN). The start codon of protein-coding genes were typical ATN (ATA, ATT, ATC, ATG) except that several special start codon (TTG, AAC, AAT, GTG) existed in *Nad1*, *COI*, *ATP8* in some species. The termination codon of all the protein coding genes was common TAR (TAA, TAG), T or TA. In the phylogenetic analysis, 6 subfamilies were clustered into separate branches, and *T. gebleri* was clustered together with *Batocera lineolata*. **[Conclusion]** *T. gebleri* has similar mitochondrial genome to those previously reported longicorn beetles. Except that several tRNA genes are rearranged, the mitochondrial genes in Cerambycidae are relatively stable. Phylogenetic reconstruction supported the monophyly of the six subfamilies and presented that *T. gebleri* was genetically close to *B. lineolata* in this study.

Key words *Thyestilla gebleri*, Cerambycidae, mitochondrial genome, phylogeny

线粒体是一种存在于各类真核生物细胞、自身拥有独立的遗传物质和遗传体系的细胞器。线粒体基因组由于普遍为母系遗传、基因组成稳定、基因排列相对保守、极少发生重组等特征,被广泛应用于群体遗传学、系统发生学、分子进化学等方面的研究(Wolstenholme, 1992; Nelson *et al.*, 2012; Driscoll *et al.*, 2015; Feng *et al.*, 2015)。昆虫作为地球上种类最多、生物学习性最为复杂的生物类群,一直以来都是线粒体基因组研究的热点对象(Boore, 1999)。天牛科 Cerambycidae 是昆虫纲 Insecta 鞘翅目 Coleoptera 中种类较多的一个科,全世界记录 3 万余种(Ślipiński *et al.*, 2011),中国记录已有 3 500 余种(王直诚和华立中, 2009),很多种类蛀食各类农作物、林木、果树及园林植物,是一类经济意义十分重要的昆虫(蒋书楠等, 1985)。截至 2016 年 12 月 20 日,GenBank 收录天牛科昆虫完整的线粒体基因组仅有 9 种,接近完整的线粒体基因组 7 种(<http://www.ncbi.nlm.nih.gov/>)。可见,相对于该科昆虫较为庞大的种类数量,现有的线粒体基因组数据还严重匮乏。

麻竖毛天牛 *Thyestilla gebleri* (Faldermann, 1835) 属于鞘翅目 Coleoptera 天牛科 Cerambycidae 沟胫天牛亚科 Lamiinae, 分布于俄罗斯东部、蒙古西北部、朝鲜、日本以及中国等广大地区。在

我国主要分布在北京、河北、河南、山东、山西、陕西、江苏等地(蒋书楠等, 1985)。主要危害桑 *Morus alba*、大麻 *Cannabis sativa*、苘麻 *Abutilon theophrasti*、苕麻 *Boehmeria nivea*、棉花 *Gossypium* spp. 等植物,以幼虫蛀食茎部或成虫取食叶、嫩梢危害,严重影响作物的产量和品质(McPartland and Hillig, 2007)。此外,该昆虫还被发现是一种重要的传粉昆虫(冯立超和王忠武, 2014)。目前关于麻竖毛天牛的研究甚少,只有少量关于其形态学、生物生态学习性和防治方法的研究报道(杨保舒, 1959; 刘时第和杨云芳, 1965; McPartland and Hillig, 2007; 宋雅琴和陈力, 2008)。本研究首次对麻竖毛天牛线粒体全基因组进行了测定,结合 GenBank 中已收录的 16 种天牛的线粒体基因组序列进行比较分析,并基于线粒体基因组数据初步探讨了 6 亚科之间的系统发育关系。该研究结果将为麻竖毛天牛的种群遗传学、分子生态学以及天牛科的分子系统发育研究奠定基础。

1 材料与方法

1.1 样品采集、鉴定和 DNA 提取

麻竖毛天牛成虫于 2015 年 6 月采自北京松山国家森林公园(采集人杨锦),采集后用 100%

无水乙醇浸泡并置于 - 20℃ 保存。试虫样品由中国科学院动物研究所林美英博士进行形态鉴定。取胸部肌肉研磨, 采用血液/细胞/组织基因组 DNA 提取试剂盒(TIANGEN)提取基因组 DNA , 用 NanoDrop2000 超微量分光光度计进行 DNA 浓度和纯度检测后, - 20℃ 保存备用。

1.2 PCR 扩增及测序

参照 Simon 等 (2006) 的引物序列, 根据已知的天牛线粒体基因组序列, 用 Primer Premier 5.0 软件设计 5 对天牛科线粒体基因组通用引物, 依据通用引物扩增得到的麻竖毛天牛序列设计 4 对特异性引物 ,引物信息见表 1。根据 TaKaRa LA Taq 酶的说明书建立 PCR 反应体系 (20 μL): 模板 0.8 μL , 正反向引物各 1 μL , 镁离子、dNTP、10×buffer 各 2 μL ,LA Taq 酶 0.2 μL , 无菌水 11 μL。PCR 扩增条件为: 94℃ 预变性 5 min ; 94℃ 变性

30 s , 45~55℃ 退火 30 s , 72℃ 延伸 1~4 min , 30 个循环 ;72℃ 彻底延伸 10 min ;10℃ 保存。其中, 退火温度根据引物设定 ,延伸时间按照扩增片段长度设定 (60 s/kb)。PCR 扩增产物用 1% 的琼脂糖凝胶电泳进行检测 ,PCR 产物测序由北京诺赛基因组研究中心完成。

1.3 序列拼接和基因注释

使用 SeqMan (Swindell and Plasterer , 1997) 对成功测序的麻竖毛天牛线粒体基因重叠片段进行拼接 ,拼接完成后提交至 MITOS Web Server 在线注释网站 (<http://mitos.bioinf.uni-leipzig.de/index.py>) , 用 Chormas pro (李雪娟等 , 2013) 寻找开放阅读框。参考 GenBank 中已知的天牛线粒体基因组序列的注释结果, 在 MEGA 5.0 (Tamura *et al.* , 2011) 中结合网站注释结果进行比对, 依次对 13 个蛋白编码基因 (PCGs)

表 1 麻竖毛天牛线粒体全基因组 PCR 扩增引物
Table 1 PCR primers for *Thyestilla gebleri* complete mitochondrial genome

引物名 Primer name	序列 5'–3' Sequence 5'–3'	目的区域 Target region
Cer-TI-F34	GCCTGAKHAAAGGRTRATTTGATA	
Cer-C1-N2353	GCTCGRGATATCWACATCTATWCC	<i>Nad2-COI</i>
TG-C1-F1455	GAGCAGGAATAGTAGGAAC	
TG-TK-N3727	TTAGAAGTAGGCGTGAGA	<i>COI-COII</i>
TG-C1-F2877	CCCCTGCTGAACATAG	
TG-A6-N4508	GATGGTCCAGTATTCCC	<i>COI-ATP6</i>
Cer-TK-F3790	CATTAGRAGRCTGAAAGYAAGYA	
Cer-TF-N6375	AGAGYDTRATATTGAAGATATYARGG	<i>ATP8-Nad3</i>
Cer-TN-F6172	AGMGGTWTRTYACTGTTAATGA	
Cer-N4-N9153	TGRGGNTATCARCCNGARCG	<i>Nad5-Nad4</i>
TG-N4-F8648	CAATAAACGCATAAGTCC	
TG-N4L-N9429	GTTTGTGAGGGGTCGT	<i>Nad4</i>
Cer-N4-F9172	CGYTCNGGYTGATANCCYCA	
Cer-LR-N13000	TTACCTYAGGGATAACAGCGTAA	<i>Nad4-rrnL</i>
Cer-CB-F11335	CAYATYCARCCHGAATGATA	
Cer-SR-N14220	AWATGYACATATCGCCCGTC	<i>Cob-rrnS</i>
TG-RL-F12912	ATAGGGTCTTCTCGTCTTTT	
TG-N2-N228	GTTCTCTGATTATCGTTGTA	<i>rrnS-Nad2</i>

以 TG 开头的引物为麻竖毛天牛特异性引物。
The primers starting with TG are specific primers of *T. gebleri*.

22 个转运 RNA (tRNA) 基因、2 个核糖体 RNA (rRNA) 基因和 A+T 富含区 (D-Loop) 进行注释。通过 tRNAscan-SE Search Server (<http://lowelab.ucsc.edu/tRNAscan-SE/>) 在线对 tRNA 基因二级结构进行预测, 并进行人工校正。注释完成后将麻竖毛天牛的线粒体基因组全序列提交 GenBank, 获得序列号。为验证注释的准确性和一致性, 对 GenBank 中已收录的 16 种天牛线粒体基因组序列在 MEGA 5.0 中进行重新注释。注释过程中, 对麻竖毛天牛和已有的天牛线粒体基因组序列的长度、A+T 含量、基因排列顺序、基因间碱基数、密码子的使用和 tRNA 结构等特征进行比较分析。已知的天牛线粒体基因组信息见表 2。

1.4 系统发育树构建

本研究从与天牛科亲缘关系较近的叶甲科中选择 3 种作为外群, 以麻竖毛天牛和 GenBank 中 16 种天牛的线粒体蛋白编码基因序列 (表 1) 共同构建系统发育树。由于 GenBank 中 *Lamiinae* sp. 等 7 种天牛的线粒体基因组蛋白编码基因 *COI* 和/或 *Nad2* 序列缺失, 本研究选取 17 种天牛共有的 11 个蛋白编码基因序列用于构建系统发育树。将这 11 个蛋白编码基因在 MEGA 5.0 中对齐, 拼接成联合蛋白序列 (共 8 531 bp) 后, 应用 MEGA 5.0 进行系统发育分析, Bootstrap Replications 设为 1 000, 采用最大似然法 (Guindon *et al.*, 2003) 构建系统发育树。

表 2 用于系统发育分析的物种信息
Table 2 Information of species used for phylogenetic analysis

亚科 Subfamily	种名 Species	序列号 Accession No.	参考文献 References
沟胫天牛亚科 Lamiinae	麻竖毛天牛 <i>Thyestilla gebleri</i>	KY292221	本研究
	云斑白条天牛 <i>Batocera lineolata</i>	NC022671	Wang <i>et al.</i> , 2012
	黄星桑天牛 <i>Psacotha hilaris</i>	NC013070	Kim <i>et al.</i> , 2009
	松墨天牛 <i>Monochamus alternatus</i>	NC024652	Li <i>et al.</i> , 2016a
	星天牛 <i>Anoplophora chinensis</i>	NC029230	Li <i>et al.</i> , 2016b
	光肩星天牛 <i>Anoplophora glabripennis</i>	NC008221	Fang <i>et al.</i> , 2016
	* <i>Lamiinae</i> sp.	JX412741	Timmermans <i>et al.</i> , 2016
天牛亚科 Cerambycinae	咖啡脊虎天牛 <i>Xylotrechus grayii</i>	NC030782	Guo <i>et al.</i> , 2016
	<i>Obrium</i> sp.	KT945156	Song <i>et al.</i> , 2015
	雾社血斑天牛 <i>Aeolesthes oenochrous</i>	NC025243	Chiu <i>et al.</i> , 2016
	栗山天牛 <i>Massicus radei</i>	NC023937	Wang <i>et al.</i> , 2014
	* <i>Closteromerus claviger</i>	HQ232804	Timmermans <i>et al.</i> , 2010
花天牛亚科 Lepturinae	* <i>Stenurella nigra</i>	JX220998	Timmermans <i>et al.</i> , 2016
	*密毛皮花天牛 <i>Rhagium mordax</i>	JX412743	Timmermans <i>et al.</i> , 2016
瘦天牛亚科 Disteniinae	* <i>Disteniazteca fimbriata</i>	JX221000	Timmermans <i>et al.</i> , 2016
膜花天牛亚 Necydalidinae	* <i>Necydalis ulmi</i>	JX220989	Timmermans <i>et al.</i> , 2016
暗天牛亚科 Vesperinae	*长翅暗天牛 <i>Vesperus conicicollis</i>	JX220996	Timmermans <i>et al.</i> , 2016
跳甲亚科 Alticinae	<i>Agasicles hygrophila</i>	KR494279	
萤叶甲亚科 Galerucinae	<i>Diabrotica barberi</i>	NC022935	
	<i>Galeruca daurica</i>	NC027114	

*表示该种的线粒体基因组序列不完整。表 4 同。

Species with a * before means its mitochondrial genome incomplete in GenBank. The same as Table 4.

2 结果与分析

2.1 麻竖毛天牛线粒体基因组基本结构

本研究获得了麻竖毛天牛完整的线粒体基因组序列,全长 15 505 bp ,A+T 含量为 74.07% ,共包括 2 个 rRNA 基因: *rrnS* (12S RNA)、*rrnL* (16S RNA);22 个 tRNA 基因;13 个蛋白编码基因: *ATP6*、*ATP8*、*COI-III*、*Nad1-6*、*Nad4L*、*Cob* ;1 个非编码区: *D-loop*(表 3)。37 个编码基因中,

14 个 (*trn-Q* , *C* , *Y* , *F* , *H* , *P* , *L1* , *V* ; *Nad1* , *Nad4* , *Nad4L* , *Nad5* ; *rrnL* , *rrnS*) 在 N 链编码 ,其余 23 个在 F 链编码 (图 1)。麻竖毛天牛基因排列顺序与已知的天牛线粒体全基因组相同 ,未发现基因重排。对比 17 种天牛线粒体基因的排列顺序 ,只有长翅暗天牛的一个 tRNA 基因发生了移位 (*trnP* 从 *trnT* 和 *Nad6* 之间移位到 *Nad6* 和 *Cob* 之间) ,其他天牛线粒体基因排列相同 ,未发现基因重排现象。

表 3 麻竖毛天牛线粒体基因组结构
Table 3 Organization of the *Thyestilla gebleri* mitochondrial genome

基因 Gene	链 Chain	位置 Location	长度 Length	反密码子 Anti codon	密码子 Codon		基因间 碱基 IGS
					起始 Start	终止 Stop	
<i>trnI</i>	F	1-64	64	GAT			0
<i>trnQ</i>	N	62-130	69	TTG			- 3
<i>trnM</i>	F	130-198	69	CAT			- 1
<i>Nad2</i>	F	199-1206	1 008		ATC	TAA	0
<i>trnW</i>	F	1205-1270	66	TCA			- 2
<i>trnC</i>	N	1263-1325	63	GCA			- 8
<i>trnY</i>	N	1326-1389	64	GTA			0
<i>COI</i>	F	1382-2924	1 543		ATT	T	- 8
<i>trnL2(UUR)</i>	F	2925-2989	65	TAA			0
<i>COII</i>	F	2990-3677	688		ATT	T	0
<i>trnK</i>	F	3678-3747	70	TTT			0
<i>trnD</i>	F	3752-3817	66	GTC			4
<i>ATP8</i>	F	3818-3973	156		ATT	TAA	0
<i>ATP6</i>	F	3967-4641	675		ATG	TAA	- 7
<i>COIII</i>	F	4641-5429	789		ATG	TAA	- 1
<i>trnG</i>	F	5432-5500	69	TCC			2
<i>Nad3</i>	F	5501-5852	352		ATT	T	0
<i>trnA</i>	F	5853-5917	65	TGC			0
<i>trnR</i>	F	5918-5979	62	TCG			0
<i>trnN</i>	F	5979-6042	64	GTT			- 1
<i>trnS1(AGN)</i>	F	6043-6109	67	TCT			0
<i>trnE</i>	F	6110-6172	63	TTC			0
<i>trnF</i>	N	6172-6236	65	GAA			- 1
<i>Nad5</i>	N	6237-7953	1 717		ATA	T	0
<i>trnH</i>	N	7951-8013	63	GTG			- 3

续表 (Table 3 continued)

基因 Gene	链 Chain	位置 Location	长度 Length	反密码子 Anti codon	密码子 Codon		基因间 碱基 IGS
					起始 Start	终止 Stop	
<i>Nad4</i>	N	8014-9340	1 327		ATA	T	0
<i>Nad4L</i>	N	9337-9624	288		ATG	TAA	- 4
<i>trnT</i>	F	9627-9689	63	TGT			2
<i>trnP</i>	N	9690-9755	66	TGG			0
<i>Nad6</i>	F	9758-10261	504		ATT	TAA	2
<i>Cob</i>	F	10261-11400	1 140		ATG	TAG	- 1
<i>trnS2</i> (UCN)	F	11399-11466	68	TGA			- 2
<i>Nad1</i>	N	11487-12437	951		TTG	TAG	20
<i>trnL1</i> (CUN)	N	12439-12501	63	TAG			1
<i>rrnL</i>	N	12502-13780	1 279				0
<i>trnV</i>	N	13781-13851	71	TAC			0
<i>rrnS</i>	N	13852-14633	782				0
<i>D-loop</i>	F	14634-15505	872				

麻竖毛天牛线粒体的基因排列相对紧凑,除 A+T 富含区外,基因间隔区一共包括 31 个碱基,分布于 6 个基因对之间,其中最长的间隔区域在 *Nad1* 与 *trnS2* 之间,间隔了 20 个碱基。基因重叠区 13 个,一共包括 42 个碱基,其中最长的区域在 *COI* 与 *trnY* 之间、*trnC* 与 *trnW* 之间,均重叠 8 个碱基。对 GeBank 已有的天牛线粒体基因组序列重新进行注释发现,基因间隔或基因重叠较大的 4 个区域为 *Nad1-trnS2*、*ATP6-ATP8*、*COI-trnY* 和 *trnC-trnW*。最大间隔区均存在于 *Nad1-trnS2* 之间,间隔了 17~22 个碱基(表 4);最大重叠区是 *trnC-trnW* 和 *COI-trnY* 两个区域,均重叠 8 个碱基,但雾社血斑天牛和咖啡脊虎天牛 *COI-trnY* 之间只间隔一个碱基;17 种天牛 *ATP6-ATP8* 之间均重叠 7 个碱基。雾社血斑天牛、*Obrium* sp.、咖啡脊虎天牛在 *trnC-trnW* 区域间隔了 1~9 个碱基,其余天牛均显示为基因重叠。

2.2 蛋白编码基因

麻竖毛天牛线粒体蛋白编码基因共有 13 个,其中 9 个(*COI-III*, *ATP6*, *ATP8*, *Nad2-3*, *Nad6*, *Cob*) 在 F 链编码,4 个(*Nad1*, *Nad4-5*, *Nad4L*) 在 N 链编码。13 个蛋白编码基因总长

度为 11 138 bp,共包含密码子 3 716 个,A+T 含量 72.62%,低于其他基因的 A+T 含量。麻竖毛天牛线粒体蛋白编码基因中,除 *Nad1* 以 TTG 为起始密码子,其余 12 个蛋白编码基因均以传统的 ATN (ATA, ATT, ATG) 为起始密码子。终止密码子除传统的 TAR (TAA, TAG) 外,*Nad3*~*Nad5*、*COI*、*COII* 以单独的 T 结尾。通过对现有的 10 种完整的线粒体基因组各区域的序列进行对比分析发现,天牛线粒体蛋白编码基因总长度变化不大(11 104~11 155 bp),但 A+T 含量有较大差异(65.86%~77.70%),且蛋白编码基因的 A+T 含量在所有线粒体基因中均为最低(表 5)。所有天牛蛋白编码基因终止密码子均为 TAR 或单独的 T 和 TA,绝大部分基因均以 ATN 为起始密码子,但少数天牛在 *Nad1*、*COI* 和 *ATP8* 中存在特殊情况:*Nad1* 基因常见以 TTG 作为起始密码子,咖啡脊虎天牛和雾社血斑天牛 *COI* 基因的起始密码子分别为 AAC 和 AAT,*Obrium* sp. 的 *ATP8* 基因以 GTG 为起始密码子。

2.3 tRNA 与 rRNA 基因

麻竖毛天牛线粒体基因组 22 个 tRNA 中,14 个在 F 链编码,8 个在 N 链编码,长度在 62 bp

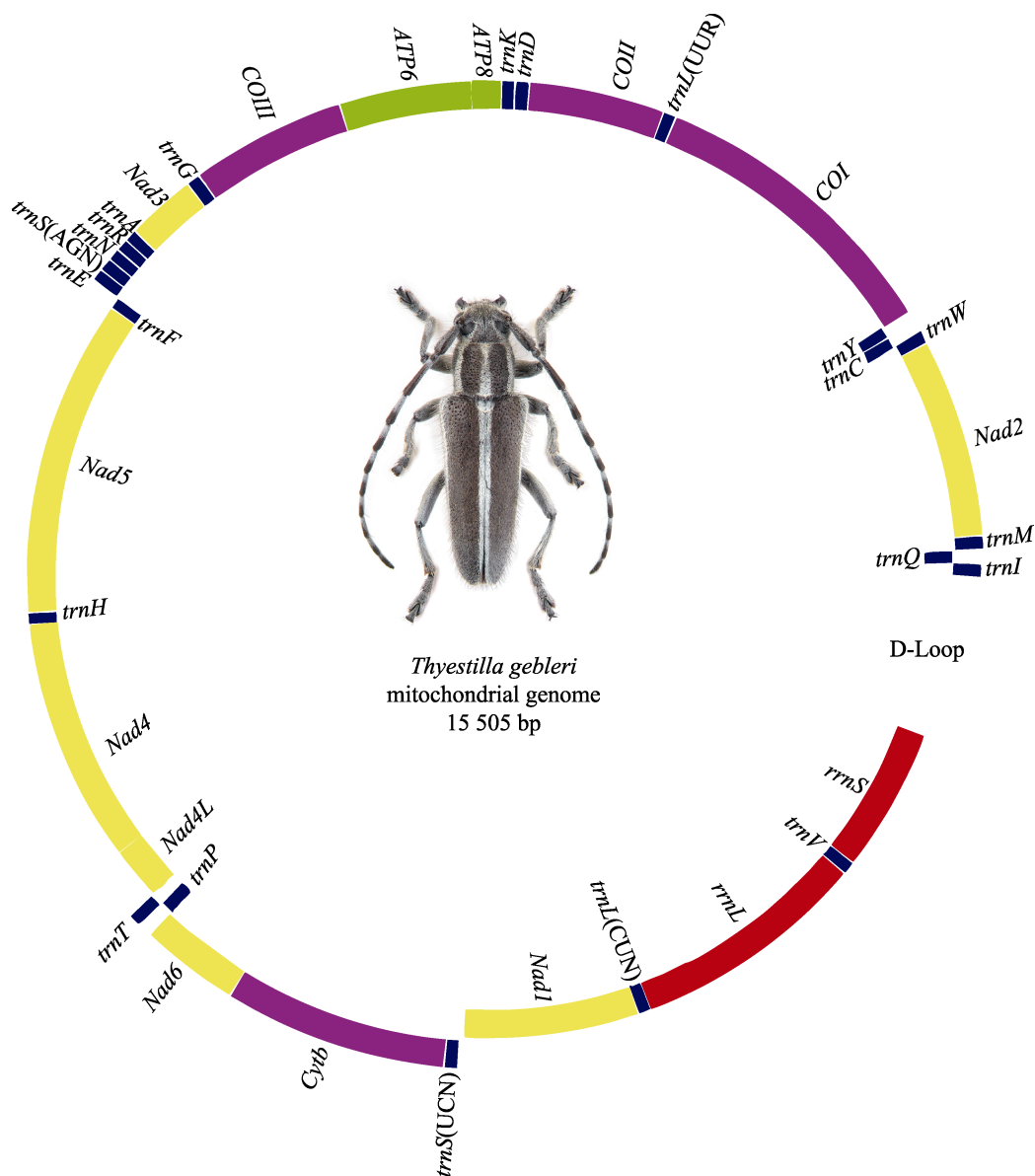


图 1 麻竖毛天牛线粒体全基因组结构图

Fig. 1 The gene map for mitochondrial genome of *Thyestilla gebleri*

麻竖毛天牛线粒体基因圈图, 圈外代表 F 链, 圈内代表 N 链。其中, 黄色: NADH 还原酶基因; 蓝色: 转运 RNA 基因; 红色: 核糖体 RNA 基因; 绿色: ATP 合成酶基因; 紫色: 细胞色素基因。

The circle diagram of *Thyestilla gebleri* mitogenome, the outer ring represents the F chain and the inner circle represents the N chain. Among them, yellow: NADH dehydrogenase genes; blue: tRNA genes; red: rRNA genes; green: ATP synthase; purple: Cytochrome gene.

(*trnR*) 到 71 bp(*trnV*) 之间。其中 *trnS1*(AGN) 和 *trnS2*(UCN) 都对应色氨酸, *trnL1*(CUN) 和 *trnL2*(UUR) 都对应亮氨酸, 其余 18 种均分别对应一种氨基酸。结合 tRNAscan-SE Search Server 对 17 种天牛的 tRNA 在线预测结果发现, 除 *trnS1*(AGN) 外, 其余 tRNA 均具有典型的三叶草结构。*trnS1*(AGN) 的 D 臂缺失, 由 8

个连续的碱基构成。

麻竖毛天牛线粒体基因组的 16S rRNA 和 12S rRNA 位于 *trnL1*(CUN) 和 *D-Loop* 区之间, 被 *trnV* 隔开。16S rRNA 长 1 279 bp, A+T 含量为 76.78%, 12S rRNA 长 782 bp, A+T 含量为 75.32%。天牛线粒体 16S rRNA 和 12S rRNA 两个区域片段差异很小, 分别为 1 271~1 288 bp 和

表 4 17 种天牛的线粒体基因组总体特征和基因间碱基数
Table 4 The characteristics of mitochondrial genomes and the ICGS of 17 species in Cerambycidae

种名 Species	长度 (bp) Length	A+T 含量 (%) A+T content	基因间碱基 Intergenic nucleotides			
			<i>Nad1-trnS2</i>	<i>ATP6-ATP8</i>	<i>COI-trnY</i>	<i>trnC-trnW</i>
<i>T. gebleri</i>	15 505	74.07	20	- 7	- 8	- 8
<i>B. lineolata</i>	15 418	79.02	20	- 7	- 8	- 8
<i>P. hilaris</i>	15 856	76.63	17	- 7	- 8	- 8
<i>Mo. alternatus</i>	15 874	71.19	17	- 7	- 8	- 8
<i>An. chinensis</i>	15 805	77.65	17	- 7	- 8	- 8
<i>An. glabripennis</i>	15 774	78.31	20	- 7	- 8	- 8
<i>X. grayii</i>	15 540	75.29	18	- 7	1	1
<i>Obrium</i> sp.	15 680	68.12	17	- 7	- 8	7
<i>Ae. oenochrous</i>	15 747	74.48	22	- 7	1	9
<i>Ma. radei</i>	15 858	71.78	20	- 7	- 8	- 8
* <i>Lamiinae</i> sp.	9 769	76.81	17	- 7		
* <i>C. claviger</i>	12 185	74.81	17	- 7		
* <i>S. nigra</i>	10 617	77.52	22	- 7		
* <i>R. mordax</i>	10 676	77.90	18	- 7		
* <i>D. fimbriata</i>	9 776	74.91	17	- 7		
* <i>N. ulmi</i>	10 656	77.29	17	- 7		
* <i>V. conicicollis</i>	10 682	74.83	19	- 7		

表 5 10 种天牛线粒体全基因组各区域的序列特征
Table 5 The regional characteristics of 10 Cerambycidae mitogenomes

种名 Species	蛋白编码基因 (PCGs)		16S 核糖体 RNA <i>rrnL</i>		12S 核糖体 RNA <i>rrnS</i>		A+T 富集区 <i>D-loop</i>	
	长度 Size	AT 含量 A+T (%)	长度 Size	AT 含量 A+T (%)	长度 Size	AT 含量 A+T (%)	长度 Size	AT 含量 A+T (%)
<i>T. gebleri</i>	11 138	72.62	1 279	76.78	782	75.32	872	84.17
<i>B. lineolata</i>	11 123	73.33	1 283	75.06	812	77.46	735	85.71
<i>P. hilaris</i>	11 155	75.84	1 272	80.11	808	78.34	1 190	78.40
<i>Mo. alternatus</i>	11 144	77.70	1 271	82.14	781	80.03	1 249	86.47
<i>An. chinensis</i>	11 135	76.36	1 274	81.16	778	78.23	1 013	85.64
<i>An. glabripennis</i>	11 104	77.00	1 274	81.32	803	77.96	1 115	87.98
<i>X. grayii</i>	11 109	73.66	1 271	79.62	806	76.18	893	85.89
<i>Obrium</i> sp.	11 136	65.86	1 288	73.21	779	70.86	1 014	79.49
<i>Ae. oenochrous</i>	11 109	69.29	1 287	75.91	801	72.78	1 074	83.24
<i>Ma. radei</i>	11 147	69.83	1 283	76.07	805	74.53	1 178	81.49

778~812 bp。A+T 含量上,除了云斑白条天牛的 12S rRNA (77.46%) 比 16S rRNA (75.06%) 略高,其它天牛均是 16S rRNA 的 A+T 含量较高。

2.4 A+T 富集区

麻竖毛天牛线粒体基因组的 A+T 富集区 (*D-Loop*) 是线粒体基因组中最大的非编码区,

位于 *trnI* 和 *trnV* 之间, 长度为 872 bp, A+T 含量为 84.17%。通过对比发现, 10 种完整的天牛线粒体基因组的 A+T 富集区长度差异较大 (735~1 249 bp), 该区域 A+T 含量最高的是光肩星天牛 (87.98%), 最低的是黄星桑天牛 (78.40%)。由于其他区域的长度差别很小, 因此 A+T 富集区的长度变化是导致天牛线粒体基因组大小变化的主要因素。

2.5 系统发育分析

本研究利用已知的 17 种天牛线粒体基因组数据, 以 3 种叶甲科 Chrysomeloidea 昆虫作为外群, 基于线粒体联合蛋白编码序列 (8 531 bp) 构建了系统发育树 (图 2)。在系统发育树中, 天牛科明显分为两支 (A、B), A 分支包括花天牛亚科和膜花天牛亚科, B 分支包括沟胫天牛亚科、暗天牛亚科、瘦天牛亚科和天牛亚科。A 分支的沟胫天牛亚科单独聚为一支 (A1), 与暗天牛亚科、天牛亚科和瘦天牛亚科聚成的 A2 分支构成姊妹群。该系统发育树支持 6 个亚科均的单系性, 与传统形态学分类吻合。沟胫天牛亚科中, 隶属于墨天牛族 Monochamini 的 3 种天牛(星天

牛、光肩星天牛和松墨天牛) 聚为一支; 天牛亚科中, 隶属于天牛族 Cerambycini 的雾社血斑天牛, 栗山天牛以高置信度聚在一起。麻竖毛天牛处于沟胫天牛亚科中较基部的位, 与云斑白条天牛亲缘关系较近。

3 讨论

本研究获得的麻竖毛天牛完整的线粒体基因组序列, 是天牛科楔天牛族中第一个被测定线粒体全基因组的昆虫。结合 GenBank 公布的 9 种天牛的完整线粒体基因组和 7 种天牛接近完整的线粒体基因组数据, 对天牛科线粒体基因组的序列特征进行了对比分析和系统发育分析。该研究结果将为麻竖毛天牛的种群遗传学研究和天牛科的系统发育重建奠定基础。

扩增麻竖毛天牛线粒体全基因组共使用了 5 对通用引物、4 对特异引物, PCR 产物均显示为单一条带。目的基因产物片段较长, 测序发现 *ATP8*、*Cob* 和 *D-loop* 区域中存在少量的 PolyA 或 PolyT 结构, 但都可以通过正反向测序的方法克服。天牛线粒体基因组基因排列顺序相对保

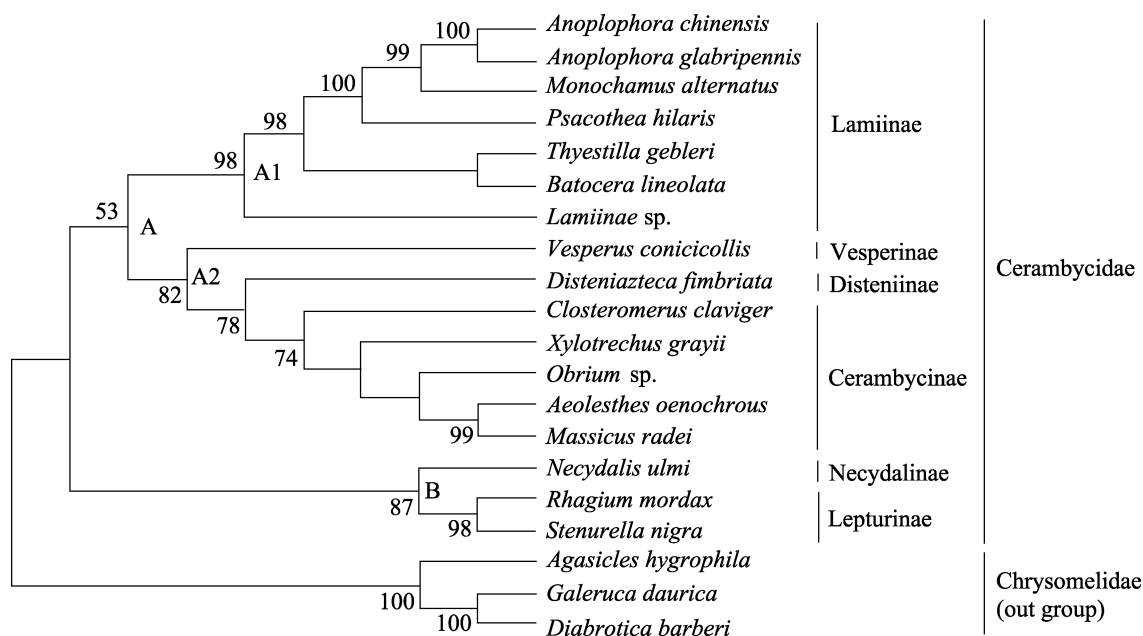


图 2 基于天牛线粒体联合蛋白序列的最大似然系统发育树

Fig. 2 Maximum Likelihood phylogenetic tree based on the concatenated PCGs

数字表示 Bootstrap 1 000 个循环的置信度 (只显示在 50% 以上的置信度)。

The numbers represent the percentage of 1 000 bootstrap replications (only above 50% showed).

守,因此本研究所设计的引物序列将为天牛科线粒体基因组的测定提供便捷和参考。

通过对 17 种天牛的线粒体基因组序列的基因排列、密码子使用和 tRNA 结构等进行比较分析发现,天牛线粒体基因排列紧凑,排列顺序相对保守,只发现长翅暗天牛 *trnP* 基因的移位。*trnP* 移位的现象在鞘翅目其他种类中也有发现 (Timmermans and Vogler, 2012)。所有已知天牛线粒体基因组在 *ATP6~ATP8* 之间都重叠 7 个碱基,在蜚蠊目、广翅目等昆虫中也存在相同的情况 (Xiao *et al.*, 2012; Zhao *et al.*, 2017),但在膜翅目 Hymenoptera 昆虫中表现不太一致 (Wei *et al.*, 2010)。

从蛋白编码区起始和终止密码子来看,天牛线粒体蛋白编码基因大部分都是以典型的 ATN 为起始密码子,只在少数种类的 *Nad1*、*ATP8*、*COI* 中存在特殊的起始密码子 (TTG、GTG、AAT 或 AAC); 终止密码子均为常见的 TAR 和不完全的 T 或 TA。不完全终止子 (T 或 TA) 在动物线粒体基因组中很常见 (Boore, 1999), Coates 等 (2005) 认为可由转录后的 3' 端的 A 补齐转化为完整的 TAA。另外,关于 *COI* 的起始密码子一直以来都存在争议,而不同的起始密码子或终止密码子会导致基因间碱基数的变化 (聂瑞娥和杨星科, 2014)。咖啡脊虎天牛和雾社血斑天牛 *COI* 基因的起始密码子之前被定义为 TTG 和 ATA (Chiu *et al.*, 2016; Guo *et al.*, 2016)。根据 Sheffield 等 (2008) 的观点,多食亚目 *COI* 的起始密码子为 AAT 或 AAC。本研究通过重新注释发现,当以 AAT 或 AAC 作为起始密码子时,咖啡脊虎天牛和雾社血斑天牛 *COI* 与 *trnY* 之间间隔很小,只间隔了 1 个碱基,更符合基因注释应该遵循的原则,因此作者认为,使用 AAC 或 AAT 作为这两种天牛 *COI* 的起始密码子更为合适。

天牛线粒体 tRNA 的二级结构除 *trnS1* (AGN) 的 D 臂缺失,其它 tRNA 基因均具有典型的三叶草结构。*trnS1* (AGN) D 臂缺失的情况在其它动物和鳞翅目 Lepidoptera、双翅目 Diptera、半翅目 Hemiptera 昆虫中也经常发现

(Wolstenholme, 1992; Liao *et al.*, 2010; Li *et al.*, 2011; Yang *et al.*, 2011)。

目前,传统的天牛分类学研究主要以形态特征为基础,但由于形态特征的复杂性和不稳定性,天牛科分类还存在诸多争议 (董抗震和杨星科, 2003)。本研究基于线粒体联合蛋白编码序列构建了天牛科 6 亚科的系统发育树,显示 6 个亚科的系统关系为 [(沟胫天牛亚科+ (暗天牛亚科+ (瘦天牛亚科+天牛亚科)) + (花天牛亚科+膜花天牛亚科)], 该结果对重建天牛科的系统发育具有一定的参考价值。在系统发育树中,麻竖毛天牛和云斑白条天牛聚为一支,但置信度较低,要进一步揭示这两种天牛的亲缘关系,还需要应用更多近缘种的线粒体基因组序列进行系统发育分析并补充相关形态和其他分子标记数据。系统发育树支持了 6 个亚科的单系性,但由于目前已知天牛的线粒体基因组的种类太少,尤其是暗天牛亚科、瘦天牛亚科和膜花天牛亚科分别只有 1 种线粒体基因组,而锯天牛亚科 Prioninae、狭胸天牛亚科 Philinae、异天牛亚科 Parandrinae、椎天牛亚科 Spondylidinae 和锯花天牛亚科 Apatophyseinae 的线粒体全基因组还是空白,因此要进一步揭示天牛科各亚科间的系统发育关系,还有待于补充大量各亚科的线粒体基因组序列信息。

致谢: 感谢林美英老师对天牛标本的形态学鉴定,感谢陈鹏燕在数据分析方面的帮助和指导,感谢窦志鹏和陈虹在系统发育树构建方面提供的帮助。

参考文献 (References)

- Boore JL, 1999. Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8): 1767–1780.
- Boore JL, 2006. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). *BMC Genomics*, 7(3): 303–309.
- Chiu WCH, Yeh WB, Chen ME, Yang MM, 2016. Complete mitochondrial genome of *Aeolesthes oenochrous* (Fairmaire) (Coleoptera: Cerambycidae): an endangered and colorful longicorn beetle. *Mitochondrial DNA*, 27(1): 686–687.

- Coates BS, Sumerford DV, Hellmich RL, Lewis LC, 2005. Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*. *International Journal of Biological Sciences*, 1(1): 13–18.
- Dong KZ, Yang XK, 2003. Progress in classification of higher taxa of the Cerambycidae. *Entomological Knowledge*, 40(3): 211–217. [董抗震, 杨星科, 2003. 天牛科高级阶元分类研究进展. 昆虫知识, 40(3):211–217.]
- Driscoll CC, Driscoll JG, Hazekamp C, Mitton JB, Wehausen JD, 2015. A tale of two markers: Population genetics of colorado rocky mountain bighorn sheep estimated from microsatellite and mitochondrial data. *Journal of Wildlife Management*, 79(5): 819–831.
- Fang J, Qian L, Xu M, Yang XJ, Wang BD, An YL, 2016. The complete nucleotide sequence of the mitochondrial genome of the Asian longicorn beetle, *Anoplophora glabripennis* (Coleoptera: Cerambycidae). *Mitochondrial DNA*, 27(5): 3299–3300.
- Feng LC, Wang ZW, 2014. Survey on the main insect-pollinator groups in Jilin city area. *Journal of Shandong Agricultural University*, 45(4): 507–514. [冯立超, 王忠武, 2014. 吉林市周边主要传粉昆虫类群构成调查. 山东农业大学学报, 45(4): 507–514.]
- Feng P, Zhao HB, Lu X, 2015. Evolution of mitochondrial DNA and its relation to basal metabolic rate. *Mitochondrial DNA*, 26(4): 566–571.
- Guindon S, Gascuel O, 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Molecular Systems Biology*, 52(5): 696–704.
- Guo K, Chen J, Xu CQ, Qiao HL, Xu R, Zhao XJ, 2016. The complete mitochondrial genome of the longicorn beetle *Xylotrechus grayii* (Coleoptera: Cerambycidae). *Mitochondrial DNA*, 27(3): 2133–2134.
- Jiang SN, Hua LZ, Pu FJ, 1985. Economic Insect Fauna of China: Coleoptera, Cerambycidae, Part 3. Beijing: Science Press. 1–205. [蒋书楠, 华立中, 蒲富基, 1985. 中国经济昆虫志: 鞘翅目, 天牛科(三). 北京: 科学出版社. 1–205.]
- Kim KG, Hong MY, Kim MJ, Im HH, Kim MI, Bae CH, Seo SJ, Lee SH, Kim I, 2009. Complete mitochondrial genome sequence of the yellow-spotted long-horned beetle *Psacotha hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among coleopteran insects. *Molecules and Cells*, 27(4): 429–441.
- Li FB, Zhang HX, Wang W, Weng HB, Meng ZQ, 2016. Complete mitochondrial genome of the Japanese pine sawyer, *Monochamus alternatus* (Coleoptera: Cerambycidae). *Mitochondrial DNA*, 27(2): 1144–1145.
- Li H, Gao J, Liu H, Liu H, Liang A, Zhou X, Cai W, 2011. The architecture and complete sequence of mitochondrial genome of an assassin bug *Agriosphodrus dohrni* (Hemiptera: Reduviidae). *International Journal of Biological Sciences*, 7(6): 792–804.
- Li W, Yang X, Qian L, An Y, Fang J, 2016b. The complete mitochondrial genome of the citrus long-horned beetle, *Anoplophora chinensis* (Coleoptera: Cerambycidae). *Mitochondrial DNA*, 27(6): 4665–4667.
- Li XJ, Yang J, Wang JH, Ren QL, Li X, Huang Y, 2013. Methods and software tools for mitochondrial genome assembly and annotation. *Chinese Bulletin of Entomology*, 50(1): 298–304. [李雪娟, 杨婧, 王俊红, 任倩俐, 李霞, 黄原, 2013. 线粒体基因组数据的分析方法和软件. 应用昆虫学报, 50(1): 298–304.]
- Liao F, Wang L, Wu S, Li YP, Zhao L, Huang GM, Niu CJ, Liu YQ, Li MG, 2010. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *International Journal of Biological Science*, 6(2): 172–186.
- Liu SD, Yang YF, 1965. Preliminary observation and control experiment of *Thyestilla gebleri*. *Acta Agriculturae Boreali Sinica*, 1(4):67–68. [刘时第, 杨云芳, 1965. 麻天牛的初步观察和防治试验. 华北农学报, 1(4): 67–68.]
- McPartland JM, Hillig KW, 2007. Longhorn beetles and botryosphaeria: A future epidemic. *Journal of Industrial Hemp*, 12(2): 123–133.
- Nelson LA, Lambkin CL, Batterham P, Wallman JF, Downton M, Whiting MF, Yeates DK, Cameron SL, 2012. Beyond barcoding: A mitochondrial genomics approach to molecular phylogenetics and diagnostics of blowflies (Diptera: Calliphoridae). *Gene*, 511(2): 131–142.
- Nie RE, Yang XK, 2014. Research progress in mitochondrial genomes of Coleoptera. *Acta Entomologica Sinica*, 57(7): 860–868. [聂瑞娥, 杨星科, 2014. 鞘翅目昆虫线粒体基因组研究进展. 昆虫学报, 57(7): 860–868.]
- Sheffield NC, Song H, Cameron SL, Whiting MF, 2008. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Molecular Biology and Evolution*, 25 (11): 2499–2509.
- Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT, 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics*, 37(1): 545–579.
- Ślipiński SA, Leschen RAB, Lawrence JF, 2011. Order Coleoptera Linnaeus, 1758. *Zootaxa*, 3148: 203–208.
- Song N, Zhang H, Yin X, Lin A, Zhai Q, 2015. The complete

- mitochondrial genome sequence from the longicorn beetle *Obrium* sp. (Coleoptera: Cerambycidae). *Mitochondrial DNA*, 29(1): 1–2.
- Song YQ, Chen L, 2008. An preliminary faunal analysis of Saperdini in China (Coleoptera: Cerambycidae: Lamiinae). *Southwest China Journal of Agricultural Sciences*, 21(2): 498–502. [宋雅琴, 陈力, 2008. 中国楔天牛族昆虫区系的初步分析(鞘翅目:天牛科:沟胫天牛亚科). *西南农业学报*, 21(2): 498–502.]
- Swindell SR, Plasterer TN, 1997. SEQMAN. Contig assembly. *Methods in Molecular Biology*, 70(1): 75–89.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28(10): 2731–2739.
- Timmermans MJTN, Dodsworth S, Culverwell CL, Bocak L, Ahrens D, Littlewood DTJ, Pons J, Vogler AP, 2010. Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. *Nucleic Acids Research*, 38(21): e197.
- Timmermans MJTN, Vogler AP, 2012. Phylogenetically informative rearrangements in mitochondrial genomes of Coleoptera, and monophyly of aquatic elateriform beetles (Dryopoidea). *Molecular Phylogenetics and Evolution*, 63(2): 229–304.
- Timmermans MJTN, Viberg C, Martin G, Hopkins K, Vogler AP, 2016. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. *Biological Journal of the Linnean Society*, 117(1): 83–95.
- Wang CY, Feng Y, Chen XM, 2012. Complete sequence and gene organization of the mitochondrial genome of *Batocera lineolata* Chevrolat (Coleoptera: Cerambycidae). *Chinese Science Bulletin*, 57(27): 3578–3585.
- Wang YT, Liu YX, Tong XL, Ren QP, Jiang GF, 2014. The complete mitochondrial genome of the longicorn beetle, *Massicus raddei*. *Mitochondrial DNA*, 27(1): 209–211.
- Wang ZC, Hua LZ, 2009. Collect and revision of name list on longicorn beetles in China. *Journal of Beihua University*, 10(2): 159–192. [王直诚, 华立中, 2009. 中国天牛名录厘定与汇总. *北华大学学报*, 10(2):159–192.]
- Wei SJ, Tang P, Zheng LH, Shi M, Chen XX, 2010. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A+T content and a long intergenic spacer between *atp8* and *atp6*. *Molecular Biology Reports*, 37(4): 1931–1942.
- Wolstenholme DR, 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology*, 141(1): 173–216.
- Xiao B, Chen AH, Zhang YY, Jiang GF, Hu CC, Zhu CD, 2012. Complete mitochondrial genomes of two cockroaches, *Blattella germanica* and *Periplaneta americana*, and the phylogenetic position of termites. *Current Genetics*, 58(2): 65–77.
- Yang BS, 1959. Observation on the life habits of *Thyestilla gebleri*. *Chinese Bulletin of Entomology*, 4(6): 187–189. [杨保舒, 1959. 麻天牛生活习性的观察, *昆虫知识*, 4(6): 187–189.]
- Yang F, Du YZ, Wang LP, Cao JM, Yu WW, 2011. The complete mitochondrial genome of the leaf miner *Liriomyza sativae* (Diptera: Agromyzidae): Great difference in the A+T-rich region compared to *Liriomyza trifolii*. *Gene*, 485(1): 7–15.
- Zhao YY, Zhang HL, Zhang YH, 2017. Complete mitochondrial genome of *Neochauliodes parasparsus* (Megaloptera: Corydalidae) with phylogenetic consideration. *Biochemical Systematics and Ecology*, 70: 192–199.