

基于线粒体 DNA 全序列 11 个蛋白编码基因 拼接序列的鳞翅目昆虫系统发育研究*

李青青¹, 段焰青², 李地艳^{1,3}, 李加敏¹, 刘晓飞⁴,
徐怀亮³, 李佛琳⁵, 周汝敏¹, 曹能¹, 叶辉^{4**}

[1. 云南师范大学 生命科学学院, 云南 昆明 650092;

2. 红云红河烟草(集团)有限责任公司技术中心, 云南 昆明 650202;

3. 四川农业大学 动物科技学院, 四川 雅安 625014; 4. 云南大学 生命科学学院, 云南 昆明 650091;

5. 云南农业大学 烟草学院, 云南 昆明 650201]

摘要: 目前已测定线粒体 DNA 全序列的鳞翅目昆虫有 12 种。由于高变异性和强核苷酸组成偏好性, 去除了两个蛋白基因和两个物种, 最终以 10 个鳞翅目物种的 11 个蛋白编码基因拼接序列对鳞翅目分子系统发育关系进行了研究。与前人的结果一致, 基于 11 个蛋白编码基因拼接序列数据所构的最大似然树支持鳞翅目, 鳞翅目内各总科和各科均为单系群。与 MINET^[1] 基于形态学数据的结果相同, 各总科的系统发育关系为 {卷蛾总科 + [螟蛾总科 + (尺蛾总科 + 蚕蛾总科)]}。蚕蛾总科内各科系统发育关系类似于 REGIER 等^[2] 的结论。

关键词: 鳞翅目; 系统发育; 线粒体 DNA 全序列

中图分类号: Q 963 文献标识码: A 文章编号: 1004-390X (2010) 02-0213-06

Phylogenetic Analysis within Lepidoptera Based on the Concatenated 11 Protein-coding Gene Sequences of Mitochondrial Genomes

LI Qing-qing¹, DUAN Yan-qing², LI Di-yan^{1,3}, LI Jia-min¹, LIU Xiao-fei⁴,
XU Huai-liang³, LI Fo-lin⁵, ZHOU Ru-Min¹, CAO Neng¹, YE Hui⁴

[1. College of Life Science, Yunnan Normal University, Kunming 650092, China;

2. Technology Center, Hongyun Honghe Tobacco (Group) Co., Ltd., Kunming 650202, China;

3. College of Animal Science, Sichuan Agricultural University, Yaan 625014, China;

4. College of Life Science, Yunnan University, Kunming 650092, China;

5. College of Tobacco Science, Yunnan Agricultural University, Kunming 650201, China]

Abstract: In this study, data set available for Lepidoptera so far, consisting of 12 complete (or almost complete) sequences of mitochondrial genomes, have been assembled. Concatenated sequences of 11 protein coding genes of 10 species were used to construct the phylogenetic relationships among Lepidoptera after removal of the high variation and extreme nucleotide composition bias. Based on these coding genes, results clearly confirmed that Lepidoptera was monophyletic by using maximum

收稿日期: 2009-04-14 修回日期: 2009-07-01

* 基金项目: 云南省自然科学基金项目 (2006C0027Q)。

作者简介: 李青青 (1976-), 女, 副教授, 博士, 主要从事昆虫分子进化的研究。

E-mail: doublelqq@yahoo.com.cn

** 通讯作者 Corresponding author: 叶辉 (1956-), 男, 教授, 博士生导师, 主要从事昆虫生态学研究。

E-mail: yehui@ynu.edu.cn

likelihood methods. Superfamilies and families were recovered, but not always strongly. Analysis of the concatenated 11 protein coding sequences by likelihood upholds the following relationships: {Tortricidea + [Pyraloidea + (Geometroidea + Bombycoidea)]}. The result wholly accords with the widely accepted phylogenetic hypotheses proposed by MINET^[1] based solely on morphological information. Within the Bombycoidea, our result reaches a similar conclusion with REGIER et al.^[2].

Key words: Lepidoptera; phylogeny; mitochondrial genomes

后生动物 (Metazoan) 的线粒体 DNA (mtDNA) 是长为 15 ~ 20 kb 的双链环状 DNA 分子^[3]。迄今已知 mtDNA 最大的昆虫是果蝇 (*Drosophila melanogaster*, NC_001709), 长 19 517 bp。后生动物的 mtDNA 编码 37 个基因, 包括 2 个核糖体 RNA (rRNA), 22 个转运 RNA (tRNA) 基因, 13 个来自 3 个氧化磷酸化复合体的蛋白质亚基编码基因^[3,4]; 唯一的非编码区是一个长度可变且富含 A, T 的区域, 在 mtDNA 的复制和转录过程中起调控和起始识别作用^[5]。在不同的后生动物中, 这些基因的排列顺序可变^[4]。

鳞翅目是昆虫纲中仅次于鞘翅目的第二大目, 分为 47 总科, 126 科及 250 亚科^[6]。目前已知约 20 万种, 中国约 8 000 种^[7]。由于具有进化快、结构简单、单拷贝和含量丰富等特点, mtDNA 已是研究鳞翅目系统学最常用最有效的分子标记^[8]。

截止目前, 鳞翅目中已测定 mtDNA 基因组全序列的物种共有 12 个, 分属于 8 个科 5 个总科: 卷蛾总科 Tortricoidea (*Adoxophyes*)、螟蛾总科 Pyraloidea (*Ostrinia*)、蚕蛾总科 Bombycoidea (*Bombyx*, *Antheraea*)、凤蝶总科 Papilionoidea (*Coreana*, *Artogeia*) 和尺蛾总科 Geometroidea (*Phthonandria*)。近年来, 一些研究者基于 mtDNA 基因组全序列来探讨鳞翅目各总科间系统发生关系^[9-12], 但所涉及的物种较少。HONG 等^[9]利用 13 个蛋白编码基因拼接序列对鳞翅目 3 个总科 7 个物种的系统发育关系进行了研究。3 个总科分别是蚕蛾总科的家蚕 (*Bombyx mori*), 野桑蚕 (*B. mandarina*), 柞蚕 (*Antheraea pernyi*) 和合目大蚕蛾 (*Caligula boisduwali*); 螟蛾总科的亚洲玉米螟 (*Ostrinia furnacalis*) 和玉米螟 (*O. nubilalis*) 以及卷蛾总科的茶小卷叶蛾 (*Adoxophyes honmai*)。除了合目大蚕蛾这个物种外, CARAPPELLI 等^[10]和 LEE 等^[11]所用的 6 个物种与 HONG 等^[9]使用的完全一样, 而在 BAE 等^[12]的

研究中, 只用了 3 个物种 (*B. mori*, *B. mandarina*, *A. pernyi*)。为了更全面地探讨鳞翅目昆虫间系统发育关系, 需要增加更多物种进行研究。

因此, 本文采用相应的分子数据分析方法和手段, 探讨迄今已测定线粒体基因组全序列的 5 总科 8 科 12 个鳞翅目昆虫间的系统发育关系, 并与其它分子序列数据^[9-14]和形态学结果进行了比较^[1,14,15]。

1 材料与方法

1.1 数据来源

12 个鳞翅目昆虫的线粒体基因组全序列数据情况见表 1。

1.2 分析方法

根据 JUNQUEIRA 等^[16]采用保守/可变位点比率 (conserved/variable sites ratio, C/V) 用来评估 13 个蛋白质编码基因的可信度。13 个蛋白质编码基因的 C/V 值结果见图 1。11 个基因 (COI, COII, COIII, ND1, ND2, ND3, ND4, ND4L, ND5, ATP6, Cytb) 的 C/V > 1, 将它们拼接后用于进一步的系统发育分析。另外 2 个基因, ND6 和 ATP8, 由于它们的 C/V < 1, 表明这两个蛋白编码基因的可信度不高, 因此在分析中予以剔除。

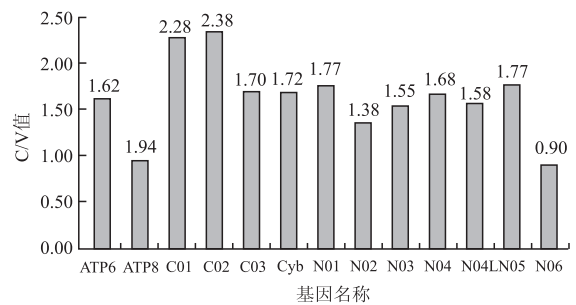


图 1 12 个鳞翅目昆虫线粒体蛋白质编码基因序列的 C/V 值
Fig. 1 The C/V ratio of aligned mitochondrial protein coding gene sequences

表 1 鳞翅目中已测定线粒体基因组全序列的昆虫

Tab. 1 The complete mitochondrial genome sequences of species in Lepidoptera

假定分类群 putatively grouping	superfamily 总科	科 family	种 species	长度/bp length	GenBank ID 号 GenBank ID	参考文献 reference
小鳞翅目, 小的蛾类 Microlepidoptera smaller moths	卷蛾总科 Tortricoidea	卷蛾科 Tortricidae	茶小卷叶蛾 <i>Adoxophyes honmai</i>	15 680	NC_008141	[11]
			亚洲玉米螟 <i>Ostrinia furnacalis</i>	14 536	NC_003368	[30]
	螟蛾总科 Pyraloidea	草螟科 Crambidae	玉米螟 <i>Ostrinia nubilalis</i>	14 535	NC_003367	[30]
大鳞翅目, 大的蝶类和蛾类 Macrolepidoptera larger butterflies and moths	蚕蛾总科 Bombycoidea	蚕蛾科 Bombycidae	家蚕 1 <i>Bombyx mori</i> 1	15 643	NC_002355	未发表 unpublished
			家蚕 2 <i>Bombyx mori</i> 2	15 656	AB 070264	[31]
			野桑蚕 <i>Bombyx mandarina</i>	15 928	NC_003395	[31]
	天蚕蛾科 Saturniidae	天蚕蛾科 Saturniidae	柞蚕 <i>Antheraea pernyi</i>	15 566	NC_004622	未发表 unpublished
			合目大蚕蛾 <i>Caligula boisduvalii</i>	15 360	NC_010613	[9]
	天蛾科 Sphingidae	天蛾科 Sphingidae	烟草天蛾 <i>Manduca sexta</i>	15 516	NC_010266	[32]
	凤蝶总科 Papilionoidea	灰蝶科 Lycaenidae	朝灰蝶 <i>Coreana raphaelis</i>	15 314	NC_007976	[33]
粉蝶科 Pieridae		菜粉蝶 <i>Artogeia melete</i>	15 140	NC_010568	未发表 unpublished	
尺蛾总科 Geometroidea	尺蛾科 Geometridae	桑尺蠖 <i>Phthonandria atrilineata</i>	15 499	NC_010522	未发表 unpublished	

用 TREE – PUZZLE 5.2^[17] 检测每一序列的碱基组成异质性 (compositional heterogeneity)。由于碱基组成的极端偏好性, 两个属于凤蝶总科的物种, 朝灰蝶 (*Coreana raphaelis*) 和菜粉蝶 (*Artogeia Melete*) 被排除在外^[9]。

11 个蛋白质编码基因的核苷酸序列用 DNASTar (DNASTar Inc.) 结合人工进行校对后拼接。鞘翅目的窗萤 (*Pyrocoelia rufa*, NC_003970) 和双翅目的果蝇 (*Drosophila yakuba*, NC_001322) 所对应的序列作为外群。最终, 包括排序中的间

隔 (gap) 在内, 长为 10 732 bp 的序列用于系统发育关系分析。

用 PAUP* 4. 0b8a^[18] 软件构建了拼接数据的最大似然树 (ML tree)。最大似然树的搜索以及对各分枝置信度 (bootstrap) 估计的参数设置如下: heuristic search, number of bootstrap iterations = 100^[19], 100 random-addition sequences, Branch-swapping algorithm: tree-bisection-reconnection branch (TBR) swapping。

用 Modeltest 3. 06^[20] 选择最大似然分析所用碱基替代模型, 以随机种子树进行搜索, 用 4 个 Markov 链模型进行 1 000 000 次循环运算。各项运算每隔 100 个循环进行 1 000 个样本抽样。每项运算最开始的 1 000 个样本被忽略 (burnin)。以显示 GTR + I + G 模型^[21] 为最适模型。核苷酸频率为 (A: 0. 414 5; C: 0. 114 4; G: 0. 087 7; T: 0. 383 4)。转换的替换速率为 (A - C: 7. 492 8; A - G: 15. 610 7; A - T: 18. 466 0; C - G: 20. 685 5; C - T: 58. 404 7; G - T: 1. 000 0。不变位点和 gamma 分布参数为 0. 186 6 和 0. 591 2。

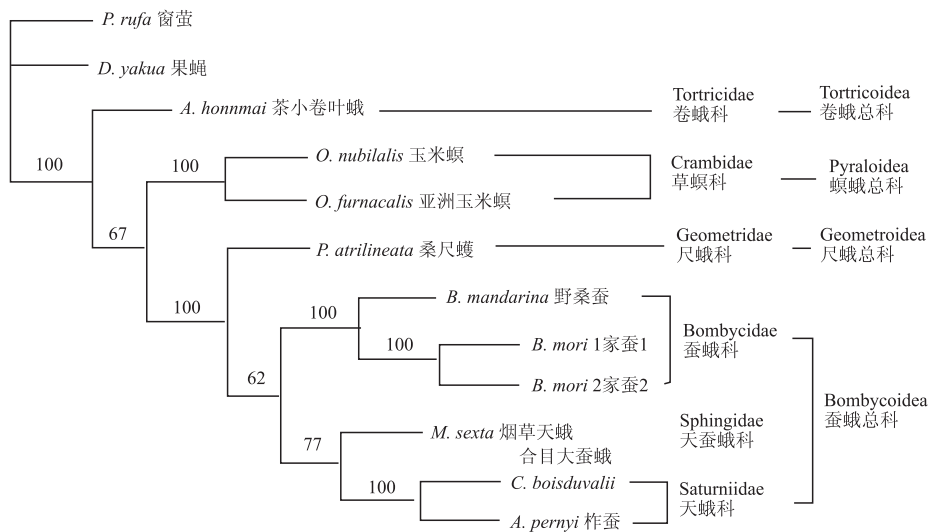
2 结果与分析

基于 11 个蛋白编码基因拼接序列所构建的 ML 系统树支持鳞翅目昆虫为单系群 (图 2), 这与前人的形态学和分子序列数据结果一致^[1,9~12,14,15]。

进化树 (图 2) 中, 12 条蛋白编码基因拼接序列包括了鳞翅目的 4 总科 (卷蛾总科, 螟蛾总科, 尺蛾总科和蚕蛾总科), 6 科 (家蚕和野桑蚕代表蚕蛾科; 柞蚕和合目大蚕蛾代表天蚕蛾科; 烟草天蛾代表天蛾科; 玉米螟和亚洲玉米螟代表草螟科; 桑尺蠖代表尺蛾科; 茶小卷叶蛾代表卷蛾科)。除蚕蛾总科作为单系群的置信度较低外 (62%), 其余总科均为单系群且置信度较高 (>80%)。

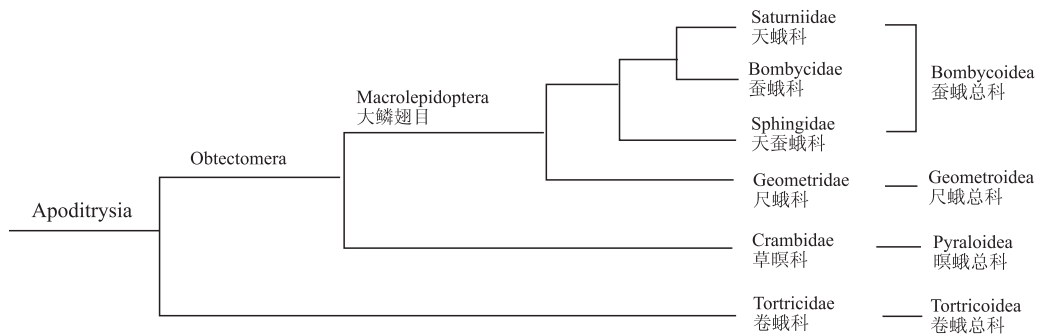
虽然置信度不高, 但总科和科均组成了各自的分支。结果支持以下的总科系统关系: {卷蛾总科 + [螟蛾总科 + (尺蛾总科 + 蚕蛾总科)]}。卷蛾总科位于所有总科的基底部, 提示卷蛾总科占据了最原始的位置。这与形态和其它分子数据结果一致。

图 2 中, 继卷蛾总科后分化出来的是螟蛾总科。尺蛾总科和蚕蛾总科互为姐妹群关系。这些结果与被人们广泛接受的 MINET^[1] 基于形态学特征得出结论一致 (图 3)。有学者将螟蛾总科和卷蛾总科归于一个假想的分类群, 即小鳞翅目 (Microlepidoptera), 而其它两个总科, 尺蛾总科和蚕蛾总科, 归于另一个假想的分类群, 大鳞翅目 (Macrolepidoptera)^[6,3,14]。形态学证据认为小鳞翅目比大鳞翅目更为原始^[6,3,14]。螟蛾总科与大鳞翅目关系较近, 共同组成 Obtectomera 后再和卷蛾总科归于 Apoditrysiian^[1] (图 3)。



枝上的数字为置信度; 树的右侧为鳞翅目科和总科。Numbers on each node indicate bootstrap values; Families and superfamilies of Lepidoptera are also indicated to the right of the tree.

图 2 基于 11 个蛋白编码基因拼接序列所构建的 ML 系统树
Fig. 2 Maximum likelihood trees based on the nucleotide sequences of the concatenated 11 protein coding genes



树的右侧为鳞翅目总科。Superfamilies of Lepidoptera are also indicated in tree.

图 3 基于MINET^[1]的形态学特征和POWELL等^[15]的幼虫取食习性而得到的鳞翅目各科系统发育关系

Fig. 3 A proposed phylogeny of the Lepidopteran families used in this study as suggested by MINET^[1] based on morphological characters and by POWELL et al.^[15] based on larval feeding habits

3 讨论

本研究中, 蚕蛾总科包括了蚕蛾科 (Bombycidae), 天蛾科 (Sphingidae) 和天蚕蛾科 (Saturniidae) 3 个科目。由于这几科的昆虫具有较高的经济价值和科研意义, 许多物种, 如家蚕 (*B. mori*) 和烟草天蛾 (*M. sexta*) 已成为生态学、生理学、和遗传学研究上的模式生物^[14,22-29]。但是, 无论是单纯基于分子数据还是基于形态学和分子数据联合分析, 蚕蛾总科内的昆虫系统发育关系上不同的研究者结论分歧较大, 甚至完全对立。

基于 5 个形态特征, MINET^[1,14] 推断 3 个科之间的关系为 [天蛾科 + (蚕蛾科 + 天蚕蛾科)], 这个结论是目前最为广泛接受的 (图 3)。

REGIER 等^[2] 利用 *period* 基因来推断鳞翅目内各科属的系统发生关系。研究结论支持蚕蛾科位于 3 个科的基部。REGIER^[22] 等收集了蚕蛾总科中的大部分昆虫样品, 扩增了 5 个核基因 (CAD, DDC, *enolase*, *period* 和 *wingless*) 长约 6570bp 的片段并进行了系统发育关系分析。虽然其所构建的系统树置信度较低 (< 50%, 他们文中的图 2), 但完全支持形态学的结果^[1,14]。

基于 CAD(680 bp) 和 EF-1a(1 240 bp) 序列数据, ZWICK^[13] 同样对蚕蛾总科中天蛾科, 蚕蛾科和天蚕蛾科昆虫的系统发育关系进行了研究, 获得了不同的进化树结果。在 MP 树中, 3 个科的位置与 REGIER 等^[2] 的结果一致, 然而在 ML 和 Bayesian 树中, 天蚕蛾科位于基部。笔者的研究结果与 REGIER 等^[2] 的相似, 所构建的似然树中, 蚕蛾科最原始, 而天蛾科和天蚕蛾科关系较

近 (图 2)。

[参考文献]

- [1] MINET J. Tentative Reconstruction of the ditrysian phylogeny (Lepidoptera, Gloassata) [J]. *Entomologica Scandinavica*, 1991, 22 (1): 69-95.
- [2] REGIER J C, FANG Q Q, MITTER C, et al.. Evolution and phylogenetic utility of the *period* gene in Lepidoptera [J]. *Molecular Biology Evolution*, 1998, 15 (9): 1172-1182.
- [3] WOLSTEN HOLME D R. Animal mitochondrial DNA: structure and evolution [J]. *International Review of Cytology - a Survey of Cell Biology*, 1992, 141: 173-216.
- [4] BOORE J L. Animal mitochondrial genomes [J]. *Nucleic Acids Reseach*, 1999, 27 (8): 1767-1780.
- [5] TAANMAN J W. The mitochondrial genome: structure, transcription, translation and replication [J]. *Biochimica et Biophysica Acta*, 1999, 1410 (2): 103-123.
- [6] KRISTENSEN N P. Lepidoptera: Moths and Butterflies [M]. 1. Evolution, Systematics, and Biogeography. Vol. IV. Berlin and New York: De Gruyter, 1999: 491.
- [7] 郑乐怡, 归鸿. 昆虫分类 (上册) [M]. 南京: 南京师范大学出版社, 1999: 805.
- [8] 李青青, 段焰青, 李佛琳, 等. 线粒体基因在鳞翅目昆虫分子系统学中的研究进展 [J]. *昆虫知识*, 2009, 46 (3): 372-382.
- [9] HONG M Y, LEE E M, JO Y H, et al.. Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects [J]. *Gene*, 2008, 413 (1-2): 49-57.
- [10] CARAPELLI A, LIO P, NARDI F, et al.. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea [J].

- BMC Evolutionary Biology, 2007, 7 (Suppl 2): S8.
- [11] LEE E S, SHIN K S, KIM M S, et al. . The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae) [J]. Gene, 2006, 373: 52 – 57.
- [12] BAE J S, KIM I, SOHN H D, et al. . The mitochondrial genome of the firefly, *Pyrocoelia rufa*; complete DNA sequence, genome organization, and phylogenetic analysis with other insects [J]. Molecular Phylogenetics and Evolution, 2004, 32 (3): 978 – 985.
- [13] ZWICK A. Molecular phylogenetics of Anthelidae and other bombycoid taxa (Lepidoptera: Bombycoidea) [J]. Systematic Entomology, 2008, 33: 190 – 209.
- [14] MINET J. The Bombycoidea: phylogeny and higher classification (Lepidoptera: Glossata) [J]. Entomologica Scandinavica, 1994, 25 (1): 63 – 88.
- [15] POWELL J, MITTER CAFANELL B. Evolution of larval food preferences in Lepidoptera [M]. KRISTENSEN N P. The Handbook of Zoology/Handbuch Der Zoologie Berlin and New York: Walter de Gruyter, 1998: 403 – 422.
- [16] JUNQUEIRA A C, LESSINGER A C, TORRES T T, et al. . The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae) [J]. Gene, 2004, 339: 7 – 15.
- [17] SCHMIDT H A, STRIMMER K VON HAESELER A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing [J]. Bioinformatics, 2002, 18 (3): 502 – 504.
- [18] SWOFFOR D L. PAUP* Phylogenetic Analysis Using Parsimony (* and other methods). Version 4.0b8a [CP]. Sunderland, MA: Sinauer Associates, 2001.
- [19] FELSENSTEIN J. Confidence limits on phylogenies: an approach using the bootstrap [J]. Evolution, 1985, 39: 783 – 791.
- [20] POSADA DCRANDALL K A. Modeltest: Testing the model of DNA substitution [J]. Bioinformatics, 1998, 14: 817 – 818.
- [21] YANG Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods [J]. Journal of Molecular Evolution, 1994, 39: 306 – 314.
- [22] REGIER J C, COOK C P, MITTER C, et al. . A phylogenetic study of the bombycoid complex (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny [J]. Systematic Entomology, 2008, 33 (1): 175 – 189.
- [23] ARUNKUMAR K P, METTA M, NAGARAJU J. Molecular phylogeny of silkmoths reveals the origin of domesticated silkmoth, *Bombyx mori* from Chinese *Bombyx mandarina* and paternal inheritance of *Antheraea proylei* mitochondrial DNA [J]. Molecular Phylogenetics and Evolution, 2006, 40 (2): 419 – 427.
- [24] REGIER J C, PAUKSTADT U, PAUKSTADT L H, et al. . Phylogenetics of eggshell morphogenesis in *Antheraea* (Lepidoptera: Saturniidae): unique origin and repeated reduction of the aeropyle crown [J]. Systematics and Biodiversity, 2005, 54 (2): 254 – 267.
- [25] LI A, ZHAO Q, TANG S, et al. . Molecular phylogeny of the domesticated silkworm, *Bombyx mori*, based on the sequences of mitochondrial cytochrome b genes [J]. J Genet, 2005, 84 (2): 137 – 142.
- [26] XIA Q, ZHOU Z, LU C, et al. . A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*) [J]. Science, 2004, 306 (5703): 1937 – 1940.
- [27] MITA K, KASAHARA M, SASAKI S, et al. . The genome sequence of silkworm, *Bombyx mori* [J]. DNA Research, 2004, 11 (1): 27 – 35.
- [28] NAGARAJU J, GOLDSMITH M R. Silkworm genomics-progress and prospects [J]. Current Science, 2002, 83 (4): 415 – 425.
- [29] WILLIS J H, WILKINS A S, GOLDSMITH M R. A brief history of Lepidoptera as model systems [M]. Molecular Model Systems in the Lepidoptera, ed. Goldsmith MRWilkinson KS. New York, NY: Cambridge University Press, 1995: 1 – 20.
- [30] COATES B S, SUMERFORD D V, HELLMICH R L, et al. . Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis* [J]. International Journal of Biological Sciences, 2005, (1): 13 – 18.
- [31] YUKUHIRO K, SEZUTSU H, ITOH M, et al. . Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina* and its close relative, the domesticated silkmoth, *Bombyx mori* [J]. Molecular Biology and Evolution, 2002, 19 (8): 1385 – 1389.
- [32] CAMERON SLWHITINHG M F. 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 (1–2): 112 – 123.
- [33] KIM I, LEE E M, SEOL KY, et al. . The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae) [J]. Insect Molecular Biology, 2006, 15: 217 – 225.