

基于线粒体 *COI* 基因全序列的直翅目部分类群 系统发育关系分析

吕红娟, 黄原*

(陕西师范大学 生命科学学院, 西安 710062)

摘要: 该研究基于直翅目 56 种昆虫的 *COI* 基因全序列构建了该目部分类群间的系统发育关系, 同时也分析了 *COI* 基因编码的氨基酸序列构建直翅目系统发育关系的可靠性。将 *COI* 序列按照密码子一、二、三位点划分, 分别计算 PBS(partioned Bremer support)值, 评估蛋白质编码基因密码子不同位点的系统发生信号强度。分析结果支持螽亚目和蝗亚目的单系性; 剑角蝗科、斑腿蝗科、斑翅蝗科、网翅蝗科和槌角蝗科 5 科均不是单系群, 科间的遗传距离在 0.107~0.153 之间变化, 与其他科相比遗传距离较小, 符合将这 5 科合并为一科(即蝗科)的分类系统, 瘤锥蝗科和锥头蝗科归为锥头蝗总科, 癩蝗科单独成为一科, 这也与 Otte (1997)系统的划分一致。根据 PBS 值的大小推断密码子第三、第一位点对系统树分支的贡献比第二位点大, 并且较长的序列含有较多的信息位点。研究也证实将各物种 *COI* 基因之间的遗传距离作为直翅目划分科级阶元的工具是可行的。

关键词: 直翅目; *COI* 基因; 系统发育

中图分类号: Q969.26; Q96; Q754 文献标志码: A 文章编号: 0254-5853-(2012)03-0319-10

Phylogenetic relationship among some groups of orthopteran based on complete sequences of the mitochondrial *COI* gene

LÜ Hong-Juan, HUANG Yuan*

(College of Life Sciences, Shaanxi Normal University, Xi'an 710062, China)

Abstract: Complete sequences of the *COI* gene were used to reconstruct the phylogenetic relationship among 56 species from Orthoptera. We also analyzed the reliability of Orthoptera phylogenetic relationship using translated amino acid sequences of the *COI* genes. The *COI* sequences were divided into three data sets on the basis of different codon positions to calculate the Partitioned Bremer support (PBS), and to test the phylogenetic signal in different codon positions of protein-coding genes. The result supports the monophyly of Caelifera and Ensifera; but the monophyly of Acrididae, Catantopidae, Oedipodidae, Arcypteridae and Gomphoceridae are not supported. The P-distances among families vary from 0.107 to 0.153, which are smaller than those of other families, being consist with the classification that these 5 families should be merged into one family (Acrididae). Chrotogonidae and Pyrgomorphidae belong to the superfamily Pyrgomorphoidea. Pamphagidae should be a family alone consistent with Otte's taxonomic system. According to the PBS values, the 3rd and 1st codon positions contribution more for the Phylogenetic tree branches than the 2nd, and longer sequences contain more informative sites. We further demonstrated that it is feasible for phylogenetic studies at family level to use the genetic distances among *COI* sequences from different species of Orthoptera.

Key words: Orthoptera; *COI* gene; Phylogeny

直翅目是不完全变态类昆虫中最大和研究最多的一个目。一般学者支持直翅目分为螽亚目 (Ensifera) 和蝗亚目 (Caelifera) 2 个亚目。目前我国普遍使用的分类系统中, 将螽亚目分成驼螽总

科、裂跗螽总科、沙螽总科、螽蜥总科、鸣螽总科、蟋蟀总科、蝼蛄总科共 7 个总科(Liu & Jin, 1992, 1993; Zheng & Gui, 1999); 将蝗亚目分为蚘总科、蝗总科、蚱总科、蚤蝼总科 4 个总科(Zheng & Gui,

收稿日期: 2011-12-19; 接受日期: 2012-03-20

基金项目: 国家自然科学基金(30970346, 31172076); 中央高校基本科研业务费专项资金资助(GK201001004)

*通信作者(Corresponding author), E-mail: yuanh@snnu.edu.cn

第一作者简介: 吕红娟, 硕士研究生, 研究方向: 动物分子生物学, E-mail: lvhongjuan@stu.snnu.edu.cn

1999), 而蝗总科下分为 9 个科, 其中在我国分布的有 8 个科(Xia, 1958; Yin, 1982)。

近年来, 线粒体 DNA(mtDNA)已成为动物分类中应用最广泛的分子标记, 例如 *COI*、*CO II*、*Cytb*、12S rRNA、16S rRNA、*ND2* 等(Liang et al, 2008; Wang & Yang, 2002; Yang & Lin, 2010)。在昆虫系统发育研究中, *COI* (cytochrome oxidase subunit I) 基因是应用频率较高的一个分子标记。*COI* 基因编码细胞色素氧化酶亚基 I, 是细胞色素 C 氧化酶的重要组成部分, 昆虫中 *COI* 基因长度在 1 500 bp 左右。Pan et al (2006) 利用 *COI* 基因的条形码对斑腿蝗科 7 种蝗虫进行了研究; Zhang (2009) 基于 28S rRNA 和 *COI* 基因对蝗总科部分种类进行了系统学研究; Wang et al (2011) 基于 18S rRNA 对直翅目主要类群的系统发育关系进行了研究; Sun et al (2006) 基于 16S rRNA 研究了中国剑角蝗科的单系性; Wang et al (2008) 基于 *Cytb* 和 *COI* 基因对蝗科中五个亚科进行了系统发育关系分析; Ding & Huang (2008) 基于 *ND2* 基因对中国斑翅蝗科部分种类进行了系统学研究。从以上专家学者的研究表

明, 对于直翅目的系统发育研究主要集中在蝗亚目, 对于螽亚目的研究较少。而 *COI* 基因作为分子标记, 主要是利用大约 600 bp 的片段鉴别亚科、科及以下物种的系统发育关系(Huang et al, 2008; Liu et al, 2010; Wang & Huang, 2006), 迄今还没有利用 *COI* 基因全序列研究直翅目内高级阶元(亚目、总科、科)系统发育关系的研究报道。

本研究基于直翅目 56 种昆虫和 3 个外群种的 *COI* 基因全序列, 旨在探讨直翅目各高级阶元之间的系统发育关系及将各物种 *COI* 基因间的遗传距离作为直翅目划分科级阶元的可行性。而本研究测出的 23 条 *COI* 序列也为直翅目昆虫起源与进化关系的研究提供了有效的数据补充。

1 材料与方法

1.1 实验标本的采集、保存与鉴定

本研究共涉及直翅目蝗亚目 40 个种, 螽亚目 16 个种, 外群 3 个种(蜚蠊目、等翅目和竹节虫目各 1 种)(表 1), 其中 23 个物种的序列由本实验室测定, 另外 33 种下载自 GenBank, 各序列来源

表 1 样品的种类、来源及 GenBank 登录号
Tab. 1 Sample sources and their GenBank accession numbers

分类地位 Classification	种 Species	采集地点 Sampling locality	GenBank 登录号 GenBank accession no.
蝗亚目 Caelifera			
蝗总科 Acridoidea			
斑腿蝗科 Catantopidae	小卵翅蝗 <i>Caryanda elegans</i>	云南勐腊县补蚌	JQ301450
	紫胫长夹蝗 <i>Choroedocus violaceipes</i>	云南普洱市	JQ301451
	长翅燕蝗 <i>Eirnephilus longipennis</i>	河北省涿鹿县杨家坪	JQ301452
	斑角蔗蝗 <i>Hieroglyphus annulicornis</i>	浙江杭州	JQ301453
	斑腿勐腊蝗 <i>Menglacris maculata</i>	云南省普洱景阳河	JQ301454
	霍山蹦蝗 <i>Sinopodisma houshana</i>	湖北武汉木兰山	JQ301455
	柯氏无翅蝗 <i>Zubovskia koeppeni</i>	黑龙江二道白河	JQ301456
	<i>Schistocerca gregaria gregaria</i>		GQ491031
	中华稻蝗 <i>Oxya chinensis</i>		EF437157
	意大利蝗 <i>Calliptamus italicus</i>		EU938373
	长翅幽蝗 <i>Ognevia longipennis</i>		EU914848
	四川凸额蝗 <i>Traulia szetschuanensis</i>		EU914849
斑翅蝗科 Oedipodidae	云斑车蝗 <i>Gastrimargus marmoratus</i>		EU513373
	亚洲飞蝗 <i>Locusta migratoria migratoria</i>		EU287446
	非洲飞蝗 <i>Locusta migratoria</i>		X80245
	东亚飞蝗 <i>Locusta migratoria manilensis</i>		GU344101
	亚洲小车蝗 <i>Oedaleus decorus asiaticus</i>		EU513374
网翅蝗科 Arcypteridae	青脊竹蝗 <i>Ceracris nigricornis nigricornis</i>	湖北神农架	JQ301457
	黑膝异爪蝗 <i>Euchorthippus fusigeniculatus</i>	河南卢氏狮子坪	JQ301458
	理塘白纹蝗 <i>Leuconemacris litangensis</i>	四川理塘	JQ301459
	隆额网翅蝗 <i>Arcyptera coreana</i>		GU324311
	中华雏蝗 <i>Chorthippus chinensis</i>		EU029161

续表

分类地位 Classification	种 Species	采集地点 Sampling locality	GenBank 登录号 GenBank accession no.
蝗亚目 Caelifera			
蝗总科 Acridoidea			
槌角蝗科 Gomphoceridae	红拟棒角蝗 <i>Gomphocerippus rufus</i>		GU294759
	李氏大足蝗 <i>Gomphocerus licenti</i>		GQ180102
	西藏大足蝗 <i>Gomphocerus sibiricus tibetanus</i>		NC_015478
剑角蝗科 Acrididae	云南卡蝗 <i>Carsula yunnana</i>	云南省宁洱	JQ301460
	威廉剑角蝗 <i>Acrida willemsei</i>		EU938372
	中华剑角蝗 <i>Acrida cinerea</i>		NC_014887
	白纹佛蝗 <i>Phlaeoba albonema</i>		EU370925
瘤锥蝗科 Chrotogonidae	乡城湄公蝗 <i>Mekongiana xiangchengensis</i>	四川乡城	JQ301461
	西藏澜沧蝗 <i>Mekongiella xizangensis</i>	西藏工布江达县城	JQ301462
	云南蝗 <i>Yunnanites coriacea</i>	云南昆明	JQ301463
癞蝗科 Pamphagidae	友谊华癞蝗 <i>Sinotmethis amicus</i>	甘肃肃州	JQ301464
	笨蝗 <i>Haplotropis brunneriana</i>	黑龙江加格达奇	JQ301465
	宽纹蠢蝗 <i>Thrinchus schrenkii</i>		NC_014610
锥头蝗科 Pyrgomorphidae	短额负蝗 <i>Atractomorpha sinensis</i>		EU263919
大腹蝗科 Pneumoridae	<i>Physemacris variolosa</i>		NC_014491
蝗亚目 Caelifera			
蚘总科 Eumastacoidea			
枕蚘科 Episactidae	郑氏比蚘 <i>Pielomastax zhengi</i>		JF411955
蝗亚目 Caelifera			
蚤蛄总科 Tridactyloidea			
蚤蛄科 Tridactylidae	日本蚤蛄 <i>Tridactylus japonicus</i>	陕西宁陕	JQ301449
	<i>Ellipes minuta</i>		NC_014488
蝻亚目 Ensifera			
蝻总科 Tettigonioidae			
蝻科 Tettigoniidae	中华寰蝻 <i>Atlanticus sinensis</i>	河北保定	JQ301444
	暗褐蝻 <i>Gampsocleis sedakovii</i>	河北涿鹿	JQ301466
	摩门蝻 <i>Anabrus simplex</i>		EF373911
	优雅蝻 <i>Gampsocleis gratioza</i>		EU527333
草蝻科 Conocephalidae	斑翅草蝻 <i>Conocephalus maculatus</i>	河北顺平	JQ301445
	疑钩额蝻 <i>Ruspolia dubia</i>		EF583824
硕蝻科 Bradyporidae	笨棘颈蝻 <i>Deracantha onos</i>		EU137664
露蝻科 Phaneropteridae	陈氏掩耳蝻 <i>Elimaea cheni</i>		GU323362
蝻亚目 Ensifera			
驼蝻总科 Rhabdiphoroidea			
驼蝻科 Rhabdiphoridae	<i>Troglophilus neglectus</i>		EU938374
蝻亚目 Ensifera			
蟋蟀总科 Grylloidea			
蟋蟀科 Gryllidae	银川油葫芦 <i>Teleogryllus infernalis</i>	河北蔚县	JQ301446
	多伊棺头蟋 <i>Loxoblemmus doenitzi</i>	河北保定	JQ301447
	黄脸油葫芦 <i>Teleogryllus emma</i>		EU557269
蛩蟋科 Encopteridae	云斑金蟋 <i>Xenorryllus marmoratus</i>	河北保定	JQ301448
蚁蟋科 Myrmecophilidae	<i>Myrmecophilus manni</i>		EU938370
蝻亚目 Ensifera			
蝼蛄总科 Gryllotalpoidea			
蝼蛄科 Gryllotalpidae	东方蝼蛄 <i>Gryllotalpa orientalis</i>		AY660929
	<i>Gryllotalpa pluvialis</i>		EU938371
蜚蠊目 Blattaria	德国小蠊 <i>Blattella germanica</i>		NC_012901
等翅目 Isoptera	<i>Reticulitermes hageni</i>		NC_009501
竹节虫目 Phasmida	海南竹节虫 <i>Ramulus hainanense</i>		NC_013185

见表 1。

1.2 基因组总 DNA 的提取

取昆虫后足股节, 用无菌水冲洗后, 放于滤纸上剪开外骨骼取出肌肉放入离心管中, 加入 300 μL 匀浆缓冲液, 剪碎并研磨, 再加入 300 μL 匀浆缓冲液。60 $^{\circ}\text{C}$ 水浴消化 5~6 h。随后用水饱和酚抽提法及氯仿-异戊醇抽提法提取基因组 DNA。再用无水乙醇及 70%乙醇提纯, 洗涤后真空干燥, 并溶解于灭菌三蒸水中, -20 $^{\circ}\text{C}$ 长期保存备用。

1.3 COI 基因的 PCR 扩增及测序

PCR 扩增 COI 基因采用的是通用引物 TY-J-1460 5'-TACAATTATCGCCTAAACTTCAGCC-3' 和 C2-N-3661 5'-CCACAAATTCAGAACATTGTCCA-3'(Simon et al, 1994)。PCR 反应体系为: 10 \times ExTaq 缓冲液 2.5 μL , 2.5 mmol/L dNTPs 1.5 μL , 25 mmol/L MgCl_2 1.5 μL , 上下游引物 (20 mmol/L)各 0.3 μL , ExTaq DNA 聚合酶 1 U, 5~10 ng/ μL , 总 DNA 模板 1.0 μL , 纯水 17.7 μL , 总反应体积 25 μL 。PCR 条件是: 95 $^{\circ}\text{C}$ 5 min 预变性; 94 $^{\circ}\text{C}$ 40 s, 52 $^{\circ}\text{C}$ 50 s, 72 $^{\circ}\text{C}$ 2 min, 共 40 个循环; 72 $^{\circ}\text{C}$ 延伸 10 min。PCR 产物用 1%的琼脂糖凝胶电泳进行检测。

PCR 产物用 0.8%琼脂糖凝胶电泳切胶回收, 用北京百泰克生物技术有限公司的离心柱型 PCR 产物试剂盒回收胶块。回收 PCR 产物在上海生工生物工程有限公司测序。测序结果使用 Staden Package (Bonfield & Staden 2002)软件包中 Pregap 4 进行拼接。

1.4 数据的处理与计算

采用 Clustal X2.0 (Larkin et al, 2007)对序列进行比对, 为了检验扩增及测序过程中可能产生的突变(包括假基因的判断), 需要将 COI 基因序列翻译成氨基酸序列。在 MEGA4.1(Kumar et al, 2008)中导入保存的 fasta 格式的序列, 指定密码子在序列中的起点, 从序列的第一位开始翻译, 将 DNA 序列翻译成氨基酸序列, 结果发现所有的序列全长都能翻译通, 证明没有产生框内终止密码子。将氨基酸序列比对后再转成 DNA 序列, 保存备用。

比对后的序列在 MEGA4.1 中进行序列组成分析, 将序列分为全数据组和根据密码子的不同位点划分的密码子第一、第二和第三位数据组共 4 个数据集, 进行碱基替换饱和分析, 分别计算碱基组成、保守位点、变异位点、转换数与颠换数的比值 (R 值)等, 并统计 COI 基因的密码子使用情况。

依据已建立的系统树, 按系统树得到的分支情况将蝗总科和螞蟥总科以科为单位划分成组, 用 MEGA 4.1 计算各组间的遗传距离。再按分支情况将蝗亚目和螞亚目以总科为单位划分成组, 计算各组间的遗传距离。

1.5 系统发育分析策略

基于完成比对的 COI 序列, 利用 Modeltest3.7 (Posada, 2003)和 MrModeltest2.2 (Nylander, 2004)选择核苷酸替换模型并计算相关参数; 利用 ProtTest2.4 (Abascal et al, 2009; Reeves, 1992)选择氨基酸替换模型并计算相关参数; 用 PAUP*4.0 b10 (Swofford, 2002)、MrBayes3.1.2 (Ronquist & Huelsenbeck, 2003)、RAxML7.0.4 (Stamatakis, 2006)和 Treefinder (Abascal et al, 2007)软件重建各物种间的系统发育关系。在整个分析过程中都选用 *Blattella germanica*、*Reticulitermes hageni* 以及 *Ramulus hainanense* 作外群。

在 PAUP*4.0 b10 中, 基于核苷酸序列用最大简约法 (maximum parsimony, MP) 构建未加权 MP 树与加权 MP 树, 均进行 1 000 次重复的自举检验。其中, 加权简约树根据 R 值大小, 采取颠换赋值 1.0, 转换赋值 0.8 的方法(Ding & Huang, 2008); 基于 K2P 距离采用邻接法(neighbor joining, NJ)构建 NJ 树; 采用最大似然法(maximum likelihood, ML), 以 Modeltest3.7 中 hLRT 筛选出的最优模型 GTR+I+G 在 PAUP*4.0 b10 中构建 ML 树, 由于 ML 树的运算量较大, 在 RAxML 中对 ML 树进行了 500 次自举检验。贝叶斯法 (Bayesian inference, BI)分析时, 以 MrModelTest 选出的最优模型 GTR+I+G, 在 MrBayes 软件中运行 4 个马尔可夫链, 共运行 150 万代, 每 100 代储存一次树, 重复一次, 在舍去老化样本后(300 代), 对剩余树构建合一树, 得到贝叶斯树。以 ProtTest2.4 筛选的氨基酸最优模型 MtREV+I+G+F 在 Treefinder 中对氨基酸序列构建 ML 树, 进行 1 000 次重复的自举检验。

运用 PAUP*4.0 b10 对各种方法构建的系统树进行了 KH (Kishino-Hasegawa) 检验和 SH (Shimodaira-Hasegawa)检验, 比较各系统树间的差异是否显著。应用 TreeRot.v3 (Sorenson & Franzosa, 2007)和 PAUP*4.0 b10 软件计算 PBS 值, 用以评估基于 COI 基因构建的系统树的每个节点的支持度。

2 结果与分析

2.1 COI 序列组成

经序列比对及手工校正, 56 条 COI 序列共有 1 566 个位点。其中有保守位点 577 个、变异位点 971 个、简约位点 783 个、自裔位点 188 个(比对中有空位)。A、T、G、C 的平均含量分别为 32.6%、34.5%、15.2%和 17.7%, A+T 含量为 67.1%, 明显高于 G+C 平均含量(32.9%), 说明 COI 的碱基组成具有偏向性。另外, 密码子不同位点的 A、T、G、C 含量也不同, 但都表现出明显的 A+T 含量偏向性, 特别在密码子第三位, A+T 含量高达 84.9%。简约信息位点(Pi)以及变异位点(V)所占比率最小的均是密码子第二位点(分别为 16.3%和 35.6%), 而最大的是密码子第三位(分别为 96.7%和 98.1%), 可见密码子第三位碱基的替换速率较快。

从 COI 基因碱基替换的统计结果可以看出, R 的平均值是 0.79, 颠换大于转换。密码子不同位点的碱基替换也表现出一定差异。密码子第三位发生碱基替换的频率明显的高于第一、二位点。这可能是由于密码子第三位承担较小的进化压力所致。

利用 MEGA4.1 统计 56 种直翅目昆虫的 COI 基因密码子使用情况, 包括密码子使用频率以及相对同义密码子使用频率(RSCU)。可以看出 NNU 和 NNA 密码子的使用频率相当高, 这说明同义密码子的使用与密码子第三位的碱基组成紧密相关。当氨基酸同时存在第三位点为 A 和 U 的同义密码子时, A 的使用频率要比 U 大的多, 如: 甘氨酸(Gly)的 GGA(2.63)和 GGU(0.92), 缬氨酸(Val)的 GUA(2.22)和 GUU(1.4)。而从 COI 基因序列组成统计也可以看出, 密码子第三位的 A+T 含量高达 84.9%, 远远高于第一、二位点, 说明密码子的摆动性主要发生在第三位。

一般直翅目昆虫线粒体基因组中 13 个蛋白质编码基因的多数基因起始密码子均为通用的 ATN 起始密码子, 但 COI 基因起始密码子呈现多样性。云南卡蝗 COI 的起始密码子为标准起始密码子 ATC, 东方蝼蛄 COI 的起始密码子为标准的 ATG, 摩门螽斯、日本蚤蝼、短额负蝗和威廉剑角蝗为 CCG, 黄脸油葫芦与疑钩额螽均为 TTA, 斑腿勐腊蝗为 ACC, 非洲飞蝗具有特殊四联起始密码子 ATTA。

COI 基因的终止密码子包括两类, 第一类是完

整的终止密码子 TAA 或 TAG; 第二类是不完整的终止密码子 T 或 TA, 例如: 疑钩额螽为 TAA, 非洲飞蝗为 TAG, 摩门螽斯为不完整的终止密码子 T。

56 条直翅目 COI 基因编码的 522 个氨基酸中, 保守位点有 207 个(39.7%), 变异位点有 308 个(59.0%), 简约信息位点有 168 个(32.2%), 自裔位点有 140(26.8%)。所有的 COI 蛋白均由 20 种氨基酸组成, 其中亮氨酸(Leu)与异亮氨酸(Ile)的含量最高, 分别为 11.86、9.74。半胱氨酸(Cys)的含量最少, 平均为 0.14, 甚至在有些物种中未出现(如陈氏掩耳螽、笨棘颈螽等), 说明 COI 的氨基酸组成存在一定的偏向性。

2.2 系统发育分析

密码子不同位点的替换饱和和分析显示, 三个位点的转换和颠换均与 P-distance 呈线性关系, 说明这些序列之间未达到饱和, 系统发育信号强, 可以用于系统发育研究。

基于核苷酸序列构建的系统发生树(NJ 树、MP 树、ML 树、BI 树)显示, 直翅目与外群完全分开, 基部都分为两支: 一支由螽亚目昆虫聚在一起形成, 另一支由蝗亚目昆虫聚在一起形成。但是这两个分支内部的系统学关系在不同系统树间存在较大差异。

加权 MP 树和 NJ 树中郑氏比螽均位于蝗亚目的基部, 蚤蝼科都位于螽亚目的基部; ML 树(图 1)和 BI 树(图 2)也包括两支, 一支是螽亚目, 一支是蝗亚目, 蚤蝼科都位于蝗亚目的基部。NJ 树、MP 树和 ML 树中, 分支末端的支持度多数大于 70%, 其他分支的支持度普遍小于 70%。

蝗亚目分支的内部差异较大, ML 树与 BI 树中相同的分支是: 蚤蝼总科位于蝗亚目基部, 依次分出的是螽总科、癞蝗科。不同分支: ML 树中瘤锥蝗科、大腹蝗科与锥头蝗科形成一个分支, 再与斑翅蝗科组成姊妹群, 在癞蝗科之后分出来; 剑角蝗科、斑腿蝗科、网翅蝗科及槌角蝗科 4 科形成一大支, 分支末端的支持度普遍较高, 但科之间的关系比较混乱, 而且分支支持度都较低。BI 树中剑角蝗科、斑腿蝗科、斑翅蝗科、网翅蝗科及槌角蝗科 5 科形成一个大支, 瘤锥蝗科、大腹蝗科与锥头蝗科形成一个分支, 再与上面 5 科形成的大支形成姊妹群, 斑腿蝗科的中华稻蝗和小卵翅蝗与剑角蝗科的中华蚱蜢和威廉剑角蝗聚在一起。

螽亚目分支内部比较一致。ML 树与 BI 树均由

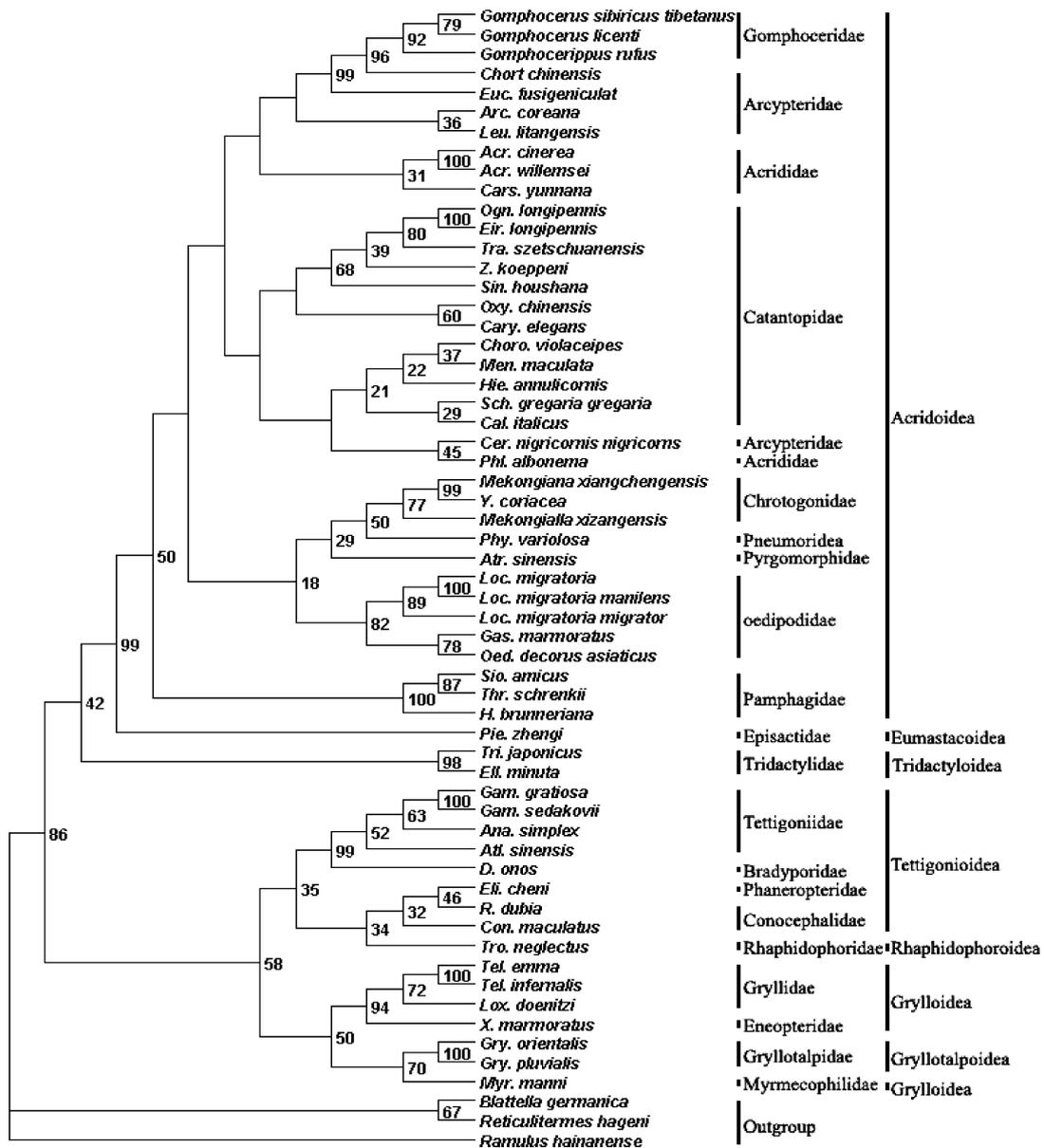


图1 PAUP*4.1 b10 基于最大似然法构建的 ML 树(数字代表自举检验值)
 Fig.1 Phylogenetic tree based on maximum likelihood method using software PAUP*4.0 b10

两大支组成: 1) 蚁蟋科的 *Myr. manni* 与蝼蛄科 2 个种形成一支, 再与 (蟋蟀科, 蛄蟋科)形成姊妹群; 2) 螽斯科 4 个种与硕螽科的 *D. onos* 形成一支, 露螽科、草螽科与驼螽科形成一支, 这两支组成一个大分支。加权 MP 树和 NJ 树中蚤蛄总科位于螽亚目基部, 但这一分支的分支支持度并不高。

通过 SH 和 KH 检验对 4 种建树方法(NJ、MP、ML、BI)构建的系统发育树进行了评估, 结果显示基于 *COI* 基因全序列构建的四种系统发育树 SH 及 KH 检验的 *P* 值均大于临界值 0.05, 即 4 种系统树

之间无显著差异。

基于氨基酸序列构建的 MP 树和 ML 树, 可以完全区分开直翅目与外群, 蝗亚目和螽亚目两个大支也很清楚, 并且支持度都较高; (蚤蛄总科, 蛄蟋总科)位于螽亚目的基部, 螽亚目内部分支清晰, 与基于核酸序列建立的 ML 中的分支一致; 蝗亚目内部分支很乱, 除个别分支与 DNA 序列建立的树一致外, 其他类群交织在一起, 不能解析总科及以下水平的系统发育关系。但是对基于氨基酸序列构建的系统树和基于核苷酸序列构建的系统树进行比较

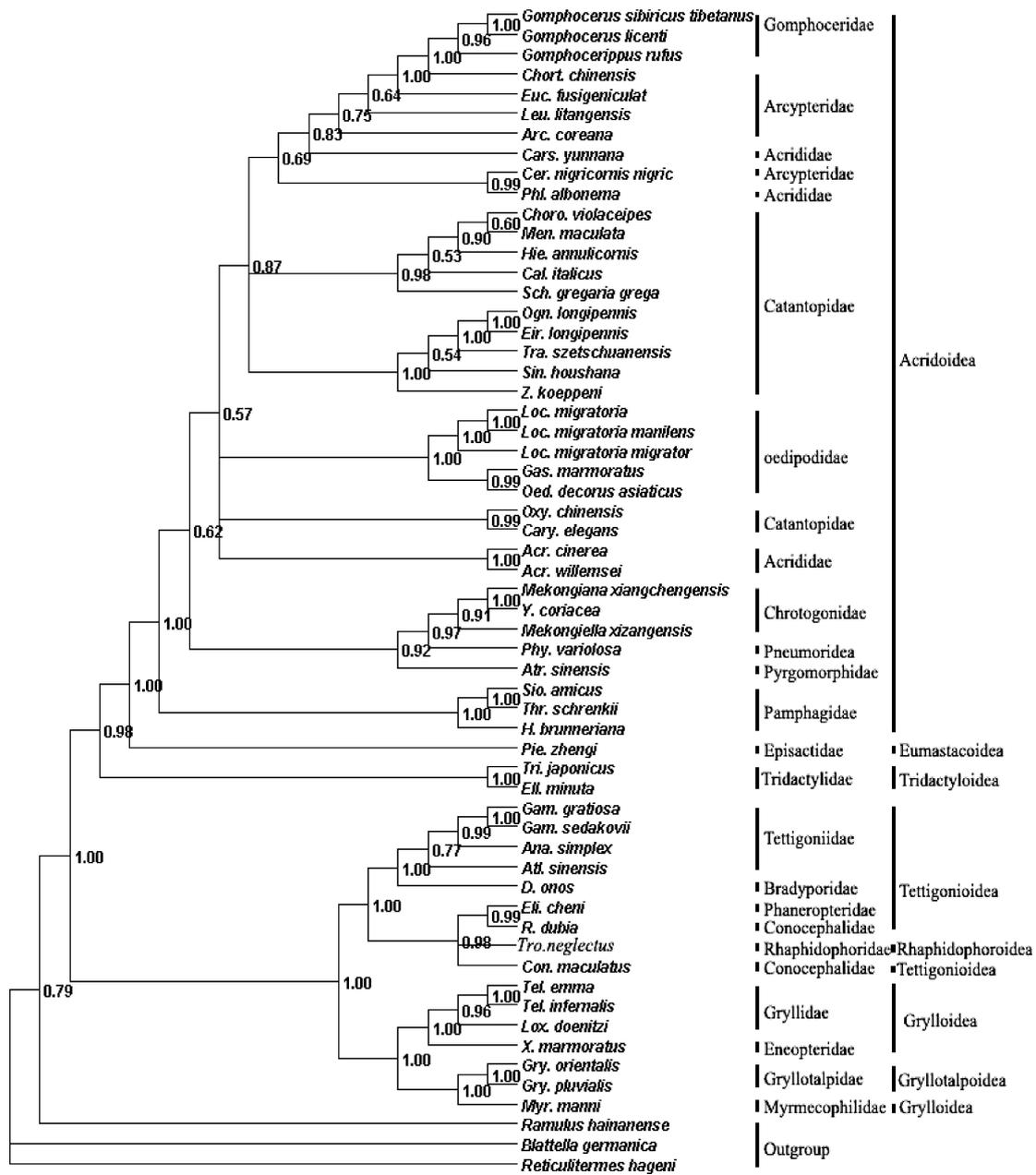


图 2 基于贝叶斯法重建的系统发育树
 Fig.2 Phylogenetic tree based on Bayesian method

可以发现：基于 DNA 序列构建的系统树末端分支的支持度普遍在 70%以上，其余分支的支持度普遍较低；基于氨基酸序列构建的系统树末端分支的支持度较高，而基部节点的支持度也普遍较高。

利用 COI 条形码的通用引物 (Pan et al, 2006), 上游引物: 5'-TYTCAACAAAYCAYAARGATATTGG- 3'; 下游引物: 5'-TAAACTTCWGGRTGWCCAAAR-AATCA-3', 分别与 59 条 COI 全长序列比对, 得到 658 bp 的序列。构建的系统发育树与 COI 全长序列构建的系统树基部的差异较大, 其末端分支大部分

支持度较高, 但是其他节点的支持度普遍比基于 COI 全长构建的系统树的支持度低。统计简约信息位点发现 COI 全长中 Pi, 占 50.0%, 而 658 bp 的片段中 Pi, 占 49.7%, 没有太大的差异。

2.3 物种间遗传距离分析

由于从系统发育树得到的各科之间及总科间分支的支持度普遍不高, 故本研究采用了组间遗传距离作为研究总科间及科间关系的参考。在 MEGA4.1 中以 P 距离计算各组之间的遗传距离(对于空位选择成对删除, 以转换加颠换计算相对遗传

距离)。

依照基于核苷酸序列构建的 ML 树的分支将系统树上的蝗总科划分为 8 个组, 螽总科分为 4 个组, 蟋蟀总科分为 3 个组。蝗总科的 8 个组为: 1 剑角蝗科、2 斑腿蝗科、3 斑翅蝗科、4 网翅蝗科、5 槌角蝗科、6 瘤锥蝗科和锥头蝗科、7 癩蝗科、8 大腹蝗科。这 8 组间的遗传距离在 0.107~0.198 之间, 瘤锥蝗科和锥头蝗科所在组与其余各组间的遗传距离最大, 均在 0.178~0.198 之间, 槌角蝗科与网翅蝗科的遗传距离最小, 为 0.107。螽总科的 4 个组为: 1 螽总科、2 硕螽科、3 露螽科、4 草螽科, 这 4 组之间的遗传距离在 0.173~0.218 之间。蟋蟀总科的 3 个组为: 1 蟋蟀科、2 蛄蟋科、3 蚁蟋科, 其中蟋蟀科与蛄蟋科的遗传距离最小, 为 0.176, 蚁蟋科与蟋蟀科、蛄蟋科的遗传距离分别为 0.216 和 0.227。

在总科级阶元之间的遗传距离方面, 还是按照 ML 树的分支将蝗亚目分为 4 个组: 1 蝗总科、2 螽总科、3 蚤蝼总科、4 大腹蝗科。这 4 组之间除蚤蝼总科外, 其他 3 组间的遗传距离在 0.189~0.232 之间, 大腹蝗科与蝗总科的距离最小, 为 0.189, 蚤蝼总科与其余各组间的遗传距离都比较大, 在 0.245~0.265 之间变化。螽亚目分为 4 个组: 1 螽总科、2 驼螽总科、3 蟋蟀总科、4 蝼蛄总科, 其中螽总科与驼螽总科的遗传距离最小, 为 0.202, 蟋蟀总科与蝼蛄总科的遗传距离是 0.211, 蟋蟀总科和蝼蛄总科与其余各组间的遗传距离均在 0.220 左右。

2.4 *COI* 基因的能效性评价

密码子第三位点的简约信息位点最多(96.7%), 密码子第一位点次之(37.0%), 由此推断 *COI* 基因序列中第三、一位点对系统树的分支贡献最大。为了进一步验证这一结论的准确性, 本研究应用 TreeRot.v3 和 PAUP4.0 b10 软件计算合一树的 PBS 值, 结果显示: 对于合一树的所有分支, 密码子第三位点的 PBS 值最大(AVG 6008.00), 第一位点次之(AVG 1699.00), 第二位点 PBS 值最小(AVG 550.00), 并且三个位点的 PBS 值差异很显著。说明密码子第三、一位点对系统树的分支贡献大, 而第二位点对系统树的分支贡献最小。这也进一步验证了 *COI* 基因序列中简约信息位点越多, 对系统树的分支贡献越大的结论。此外, 不同长度的 *COI* 基因片段计算出的 PBS 值不同, 并且 *COI* 基因的长度与

PBS 几乎呈正相关关系 (即基因序列越长, PBS 值越大), 而 *COI* 基因平均长度都在 1 500 bp 左右, 在 13 个蛋白质编码基因中仅次于 ND5(1 700 bp 左右), 这也意味着较长的基因可能含有较多的信息位点。

一般来说, 不同 DNA 片段由于其进化速率不同, 其系统发育信号不一样, 解决系统发育问题所需的最少位点数就不同。快速进化的片段对于新近分歧的类群有较强的信号, 可用较短的序列; 而对系统树基部分枝来说有很弱的信号, 需要较长的序列。并且对于大多数序列数据, 其系统发育推论的稳定性和可靠性随序列数据的增加而加强, 本研究也验证了以上说法。

从基于核苷酸序列及氨基酸序列构建的系统树的比较可以得出结论: *COI* 的核苷酸序列进化较快, 含有的变异信息更多, 更适合用来区分亚科及以下水平的系统发育关系; *COI* 的氨基酸序列进化较慢, 较为保守, 更适合区分目、亚目水平的系统发育关系。

3 讨论

3.1 不同处理方法对系统发育关系重建的影响

在进行简约法分析时, 对转换与颠换的加权和以及对密码子 3 个不同位点的加权得到的系统树是有差异的。基于氨基酸序列重建的系统发育树只能区分直翅目、蝗亚目以及螽亚目的单系性, 对于总科、科间及以下阶元均不能很好地区分。可能是由于 *COI* 氨基酸序列含有的信息较少, 只适合研究远缘物种的进化关系。PAUP*4.0 b10 和 Treefinder 软件构建的 ML 树与 MrBayes 软件构建的贝叶斯树比较一致, 均能较好的解析直翅目类群的系统发育关系。此外, 本研究为了证实同一基因的不同片段对系统发育树的影响, 我们将 *COI* 全长基因划分成不同的片段 (包括从 5'端、3'端分别取 650 bp 左右的片段和根据 *COI* 条形码的通用引物比对出的 658 bp 左右的片段), 对划分出来的 3 个片段分别建树, 然后与 *COI* 全长序列建立的系统树做比较, 发现从 5'端截取的片段和 *COI* 条形码序列构建的系统树与 *COI* 全长基因在末端分支上较一致, 而其余分支的支持度只有基于 *COI* 全长基因构建的系统树最高。

3.2 基于 *COI* 基因的直翅目各高级阶元间的系统发育关系

本研究使用 *COI* 基因全长序列重建了直翅目部分类群的系统发生树, 从不同方法构建的系统发

生树来看, 直翅目的单系性能够得到支持, 蝗亚目和螽亚目的单系性也能够得到支持。Fenn et al (2008) 利用线粒体基因组全序列数据获得的系统发育树也支持直翅目及其两个亚目的单系性。ML 树与贝叶斯树中螽亚目各类群的系统发育关系较一致, 加权 MP 树和 NJ 树与之不同的是蚤蛄总科位于螽亚目基部。蝗亚目内部的关系发育关系有较大的差异。

除蚤蛄总科外, 蝗亚目的其余物种在各系统树中均聚在一起, 构成单系群, 这些类群可以分成 3 部分, 1) 蝗总科的 *Pie. zhengi* 位于基部, 2) 瘤锥蝗科、锥头蝗科和大腹蝗科组成一支, 3) 包括癞蝗科在内的其余物种。分支中各总科和科级阶元的划分与 Otte et al (2011) 系统基本一致, 即剑角蝗科、斑腿蝗科、斑翅蝗科、网翅蝗科、槌角蝗科归到蝗科, 瘤锥蝗科和锥头蝗科归为锥头蝗总科, 癞蝗科单独成为一科。

所有的系统树中, 网翅蝗科与槌角蝗科各自都不构成单系, 但是两科的代表物种总是合并构成一个单系分枝, 支持 Otte et al (2011) 分类系统中的网翅蝗科与槌角蝗科归为一个亚科, 作为槌角蝗亚科。ML 树中斑翅蝗科代表物种聚成一个单系群, 并与瘤锥蝗科、大腹蝗科与锥头蝗科的分支组成姊妹群。在夏凯龄的分类系统中, 剑角蝗科一直被看作是单系群, 包含 6 个亚科 (Xia, 1958), 但是根据 COI 基因重建的 3 棵系统树得出的系统发生关系与夏凯龄的分类系统差别较大, 都不支持剑角蝗科是单系群, 而且近年来的多个研究 (Sun et al, 2006) 对其单系性争议较大。斑腿蝗科的单系性也不被支持。Liu & Jiang (2005) 用 18S rDNA 对蝗总科分子系统研究中, 同样发现网翅蝗科、剑角蝗科、斑腿蝗科无法形成单系群。而除 ML 树外的其余系统树中, 剑角蝗科、斑腿蝗科、网翅蝗科、斑翅蝗科、槌角蝗科 5 科共同组成一个单系群, 5 科间的遗传距离在 0.107~0.153 之间变化, 与其他科相比遗传距离较小, 支持这 5 科合并为一科, 即蝗科。这也与 Otte et al (2011) 系统的划分一致。国外普遍认同的 Otte 系统 (Otte, 1997) 从蝗总科中划分出了锥头蝗

总科和大腹蝗总科, 而在本研究中瘤锥蝗科、大腹蝗科和锥头蝗科聚在一起, 无法从蝗总科中分出来。

螽亚目在各系统发育树中均分为 2 支, 蟋蟀总科和蝼蛄科组成一支, 剩余类群聚为一支。传统的分类系统一般认为蟋蟀科和蝼蛄科有较近的进化关系, 有的分类系统将蝼蛄科作为总科, 与蟋蟀总科并列 (Zheng & Gui, 1999); 有的分类系统将蝼蛄科视为蟋蟀总科下的一个科 (Zhang et al, 2008)。结合遗传距离的计算结果, 蟋蟀总科与蝼蛄科遗传距离是 0.211, 比与其余各组间的遗传距离小, 推测两者的进化关系较近。但由于在 ML 树中这个分支的支持度不高(50.0), 所以不能完全确定蟋蟀总科和蝼蛄科的关系, 建议补充更多物种后做一步研究。

螽亚目与硕螽科 *D.onos* 聚在一起, 科间的遗传距离为 0.173, 亲缘关系较近; 草螽科与露螽科相互交织聚成一支; 螽亚目中蟋蟀科与蝼蛄科的遗传距离最小 0.176, 在系统树上也总是聚在一起, 蚁蟋科与蟋蟀科、蝼蛄科的遗传距离分别为 0.216、0.227, 且蚁蟋科与蝼蛄科总是聚在一起, 分支支持度也都较高。

本研究基于 COI 基因全序列, 应用 4 种建树方法重建了直翅目部分类群的系统发生树, 由于选取的物种有限, 并不能准确解析直翅目内部各个阶元的系统发育关系。所有系统发生树除末端分支的支持度较高外, 其余的分支支持度普遍偏低, 不足以准确解决较高阶元的系统发育关系, 但是与基于 16S rRNA、18S rRNA 建立的系统发生树有相似的分支, 因此也能够为直翅目昆虫系统发育研究提供一些有效的信息。根据系统树的分支支持度仍可以推测: COI 基因更适合于分析亲缘关系较近的亚科及以下类群的系统关系, 氨基酸序列更适合分析目及亚目高级阶元的系统发育关系。研究也证实了将各物种 COI 基因间的遗传距离作为直翅目划分子级阶元的工具具有一定的参考意义。

参考文献:

Abascal F, Posada D, Zardoya R. 2007. MtArt: A new model of amino acid replacement for Arthropoda [J]. *Mol Biol Evol*, **24**(1): 1-5.
Abascal F, Zardoya R, Posada D. 2009. ProtTest: Selection of best-fit models of protein evolution version 2.4 [J]. *Bioinformatics*, **21**(9):

2104-2105.
Bonfield JK, Staden R. 2002. ZTR: A new format for DNA sequence trace data[J]. *Bioinformatics*, **18**(1): 3-10.
Ding FM, Huang Y. 2008. Molecular evolution and phylogenetic analysis of

- some species of Oedipodidae (Orthoptera: Caelifera) in China based on complete mitochondrial ND2 gene[J]. *Acta Entomol Sin*, **51**(1): 55-60. [丁方美, 黄原. 2008. 基于线粒体 ND2 基因的中国斑翅蝗科部分种类分子系统学研究(直翅目: 蝗总科). 昆虫学报, **51**(1): 55-60.]
- Fenn JD, Song H, Cameron SL, Whiting MF. 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data [J]. *Mol Phylogenet Evol*, **49**(1): 59-68.
- Huang DW, Meier R, Todd PA, Chou LM. 2008. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding [J]. *J Mol Evol*, **66**(2): 167-174.
- Kumar S, Nei M, Dudley J, Tamura KK. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences [J]. *Brief Bioinform*, **9**(4): 299-306.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0 [J]. *Bioinformatics*, **23**(21): 2947-2948.
- Liang G, Li T, Yin ZH, Lei FM. 2008. Molecular phylogenetic analysis of some Fringillidae species based on mitochondrial CoI gene sequences[J]. *Zool Res*, **29**(15):465-475. [梁刚, 李涛, 尹祚华, 雷富民. 2008. 利用 CoI 基因序列对雀科鸟类的分子系统发育关系初探. 动物学研究, **29**(15):465-475.]
- Liu DF, Jiang GF. 2005. Molecular phylogenetic analysis of Acridoidea based on 18S rDNA with a discussion on its taxonomic system [J]. *Acta Entomol Sin*, **48**(2): 232-241. [刘殿锋, 蒋国芳. 2005. 基于 18S rDNA 的蝗总科分子系统发育关系研究及分类系统探讨. 昆虫学报, **48**(2): 232-241.]
- Liu XW, Jin XB. 1992-1993. List of Chinese Stenopelmatoidea and Terrigonoidea (Grylloptera). *Contr. Shanghai Inst Entomol*, **11**: 99-118. [刘宪伟, 金杏宝. 1992-1993. 中国螞蚱名录. 昆虫学研究集刊, **11**: 99-118.]
- Liu Y, Song Y, Li XY. 2010. DNA barcode technology in insect molecular application about mitochondrial COI gene [J]. *Plant Quarantine*, (2): 46-50. [刘勇, 宋毓, 李晓宇. 2010. 基于线粒体 COI 基因的 DNA 条形码技术在昆虫分子鉴定中的应用. 植物检疫, (2): 46-50.]
- Nylander JAA. 2004. MrModeltest v2.2. Program distributed by the author[EB/OL]. Evolutionary Biology Centre, Uppsala University.
- Otte D, Eades DC, Cigliano MM. Orthoptera Species File Online[EB/OL]. Version2.0/4.0. [2011-03-01]. <http://Orthoptera.SpeciesFile.org>.
- Pan CY, Hu J, Zhang X, Huang Y. 2006. The DNA barcoding application of mtDNA COI genes in seven species of Catantopidae (Orthoptera) [J]. *Entomotaxonomia*, **26**(2): 103-111. [潘程莹, 胡婧, 张霞, 黄原. 2006. 斑腿蝗科 Catantopidae 七种蝗虫线粒体 COI 基因的 DNA 条形码研究. 昆虫分类学报, **26**(2): 103-111.]
- Posada D. 2003. Using Modeltest and PAUP* to select a model of nucleotide substitution [J]. *Curr Protoc Bioinform*, **6**. 5. 1-6. 5. 14.
- Reeves JH. 1992. Heterogeneity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA [J]. *J Mol Evol*, **35**(1): 17-31.
- Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models [J]. *Bioinformatics*, **19**(12): 1572-1574.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers [J]. *Ann Entomol Soc Am*, **87**(6): 651-701.
- Sorenson MD, Franzosa EA. 2007. TreeRot. v3 [M]. Boston, MA: Boston University.
- Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models [J]. *Bioinformatics*, **22**(21): 2688-2690.
- Sun ZL, Jiang GF, Huo GM, Liu DF. 2006. A phylogenetic analysis of six genera of Acrididae and monophyly of Acrididae in China using 16S rDNA sequences (Orthoptera, Acridoidea) [J]. *Acta Zool Sin*, **52**(2): 302-308. [孙正莉, 蒋国芳, 霍光明, 刘殿锋. 2006. 基于 16S rDNA 序列探讨中国剑角蝗科的单系性及其六属的系统发育关系. 动物学报, **52**(2): 302-308.]
- Swofford DL. 2002. PAUP*. Phylogenetic Analysis Using Parsimony. (*and Other Methods) [EB/OL]. Version 4.0b10. Sunderland, Massachusetts: Sinauer Associates.
- Wang BX, Yang LF. 2002. Phylogenetic utilities of mitochondrial DNA sequences in the study of insect systematics[J]. *Entomol Knowl*, **39**(2): 89-91. [王备新, 杨莲芳. 2002. 线粒体 DNA 序列特点与昆虫系统学研究. 昆虫知识, **39**(2): 89-91.]
- Wang NX, Feng X, Jiang GF, Fang N, Xuan WJ. 2008. Molecular phylogenetic analysis of five subfamilies of the Acrididae (Orthoptera:Acridoidea) based on the mitochondrial cytochrome b and cytochrome c oxidase subunit I gene sequences [J]. *Acta Entomol Sin*, **51**(11): 1187-1195.
- Wang X, Huang B. 2006. Advancement of DNA Barcoding in Animal Taxonomy [J]. *Biotechnol Bull*, (4): 67-71. [王鑫, 黄兵. 2006. DNA 条形码编码技术在动物分类中的研究进展. 生物技术通报, (4): 67-71.]
- Wang XY, Zhou ZJ, Huang Y, Shi FM. 2011. The phylogenetic relationships of higher Orthopteran categories inferred from 18S rRNA gene sequences [J]. *Acta Zootaxon Sin*, **36**(3): 627-638. [汪晓阳, 周志军, 黄原, 石福明. 2011. 基于 18S rRNA 基因序列的直翅目主要类群系统发育关系研究. 动物分类学报, **36**(3): 627-638.]
- Xia KL. 1958. Taxonomic Essentials on Acrididae in China[M]. Beijing: Science Press, 1-239. [夏凯龄. 1958. 中国蝗科分类概要. 北京: 科学出版社, 1-239.]
- Yang DM, Lin Y. 2010. Application of mitochondrial DNA gene order in molecular systematics of insects [J]. *Guangdong Agric Sci*, **37**(6): 188-190. [杨德明, 林杨. 2010. 线粒体 DNA 基因序列在昆虫分子系统学研究中的应用. 广东农业科学, **37**(6): 188-190.]
- Yin XC. 1982. On the taxonomic system of Acridoidea from China [C]. *Acta Biologica Plateau Sinica*, (1): 69-99. [印象初. 1982. 中国蝗总科 (Acridoidea) 分类系统的研究. 高原生物学研究集刊, (1): 69-99.]
- Zhang HJ, Huo KK, Chen R. 2008. Comparison of Xia's taxonomic systematics with Otte's [J]. *J North West Univ: Sci Online*, **6**(4): 0358. [张宏杰, 霍科科, 陈锐. 2008. 中国夏氏与 Daniel Otte 蝗虫分类系统比较. 西北大学学报: 自然科学网络版, **6**(4): 0358.]
- Zhang XJ. 2009. Study on Molecular Phylogeny of Acridoidea Based on 28SrDNA and COI Gene [D]. S.M. thesis, Shaanxi Normal University. [张小静. 2009. 基于 28SrDNA 和 COI 基因序列的蝗总科部分种类分子系统学研究. 陕西师范大学.]
- Zheng LY, Gui H. 1999. Insect Classification [M]. Nanjing: Nanjing Normal University Press. [郑乐怡, 归鸿. 1999. 昆虫分类. 南京: 南京师范大学出版社, 254-257.]