

基于进化树、距离和特征的 DNA 条形码方法研究 ——以百花山地区草螟科为例*

杨聪慧^{1**} 金倩¹ 陈付强² 武春生² 张爱兵^{1***}

(1. 首都师范大学生命科学学院 北京 100048; 2. 中国科学院动物研究所

动物进化与系统学重点实验室 北京 100101)

摘要 DNA 条形码是一种快捷高效的分子鉴定新技术,近年来在动物分类学领域中得到迅速的发展和应用。在条形码的研究中有基于进化树、距离和特征 3 种常用的分析方法:第 1 种方法需要构建系统发育树,分析样本在树上的聚类情况;第 2 种方法依赖于物种种内和种间的序列差异;第 3 种则是通过一系列的诊断特征位点来鉴定物种。本研究扩增了北京百花山地区 14 种草螟科昆虫 88 个样本的线粒体细胞色素 C 氧化酶亚基 I (mitochondrial cytochrome c oxidase subunit I, COI) 基因片段,分别基于进化树、距离和特征方法进行了分析,以探讨不同 DNA 条形码方法在草螟科物种鉴定中的可行性。结果表明:在使用邻接法 (neighbor-joining, NJ) 构建的系统发育树上,14 种草螟物种各自聚成一个单系,均被成功区分。基于 Kimura 双参数模型计算遗传距离得出,种内和种内有一个明显的“barcoding gap”,且 ABGD 软件对样本的划分完全符合形态鉴定结果。在所有的草螟物种中都找到了诊断核苷酸位点,基于特征来鉴定草螟物种的成功率为 100%。结果显示,这 3 种方法对于本研究中所涉及的草螟都具有较好的区分,基于 COI 基因的 DNA 条形码可以作为一种有效的工具在草螟科昆虫的物种鉴定中进行应用。

关键词 鳞翅目, 草螟, DNA 条形码, COI 基因, 系统发育树, 遗传距离, 诊断特征

Application of tree-based, distance-based and character-based DNA barcoding methods for the identification of Crambidae species (Insecta: Lepidoptera) in Baihuashan, China

YANG Cong-Hui^{1**} JIN Qian¹ CHEN Fu-Qiang² WU Chun-Sheng² ZHANG Ai-Bing^{1***}

(1. College of Life Sciences, Capital Normal University, Beijing 100048, China;

2. Key Laboratory of Zoological Systematics and Evolution, Chinese Academy of Sciences, Beijing 100101, China)

Abstract DNA barcoding has recently been proposed as a promising tool for rapid and accurate species identification in a wide range of animal taxa. Three broad methods (tree-based, distance-based and character-based) are commonly used. The first method requires the recovery of species as discrete clades (monophyly) on a phylogenetic tree. The second method is based on the degree of DNA sequence variation within and between species. The last method characterizes species through a series of diagnostic characters. To explore the feasibility of DNA barcoding based on different methods in the identification of Crambid moths, the COI genes of 88 samples belonging to 14 species of the Crambidae from Baihuashan were analyzed. The results show that samples belonging to the same species cluster in one monophyly and all species were successfully distinguished by the NJ phylogenetic tree. The intraspecific and interspecific genetic distances were calculated using the Kimura-2-parameter model. An obvious barcoding gap between intraspecific and interspecific divergences for COI genes was found, while the groups automatically defined by ABGD perfectly match the groups defined by morphological characters. Diagnostic nucleotide positions were found in each species and the character-based barcode

* 资助项目:国家自然科学基金项目(31272340);北京市人才强教项目(PHR201107120)。

** E-mail: onionyc@126.com

*** 通讯作者, E-mail: zhangab2008@mail.cnu.edu.cn

收稿日期:2012-12-18, 接受日期:2012-12-28

method successfully identified every species. This study indicates that the species within the Crambidae can be readily distinguished using these three barcoding approaches based on the COI gene.

Key words Lepidoptera, Crambidae, DNA barcoding, COI gene, phylogenetic tree, genetic distance, diagnostic characters

物种鉴定和分类属于传统生物分类学范畴,人类在利用有益生物资源和防治有害生物的生产实践中,在保护生物多样性和维持全球健康稳定的生态系统过程中,都依赖于分类学家丰富的专业知识和准确的鉴定能力对物种进行正确地描绘、鉴定和管理(马英和鲁亮,2012)。但是,目前由于形态学鉴定本身的局限性和不断缩减的分类学家队伍,使分类学的发展面临巨大的挑战(肖金花等,2004)。作为一种标准、快捷、低成本的物种鉴定新方法,DNA 条形码(DNA barcoding)技术现已成为生物分类学研究的热点和进展最迅速的前沿领域之一。它利用一段标准 DNA 序列(在动物类群中通常是线粒体细胞色素 C 氧化酶亚基 I, mitochondrial cytochrome c oxidase subunit I, COI),不仅能够准确快速的区分和鉴定物种,也是发现、鉴定新种与隐存种的良好工具(Hebert *et al.*, 2003a; 2003b)。

迄今为止,基于 COI 基因的 DNA 条形码技术已在大部分动物类群的物种鉴定中被证明是行之有效的,包括昆虫(Hajibabaei *et al.*, 2006)、鸟类(Tavares and Baker, 2008)、鱼类(Ward *et al.*, 2009)、蚁类(Smith *et al.*, 2005)、甲壳类(Radulovici *et al.*, 2009)、藻类(Saunders, 2005)和原生动物(Wong and Hanner, 2008)等的鉴定。在这些运用 DNA 条形码技术进行物种识别的研究中,有 3 种方法比较常见,分别是基于进化树、距离和特征的方法。

进化树法(tree-based method)主要利用物种在系统发生上彼此互为单系群的原理(Hebert *et al.*, 2003b),通过构建分子系统发育树,分析目的序列和相邻近序列的进化关系,从最邻近参考序列来确定物种种名(马英和鲁亮,2010)。这种方法直观快速,可以在树图上更直观的展示鉴定效果。距离法(distance-based method)则是基于物种间的遗传距离来鉴定物种,它首先将目的 DNA 序列矩阵转变成遗传距离矩阵,然后基于“barcoding gap”的概念(Hebert *et al.*, 2004; Burns

et al., 2007),即种内和种间序列差异的一个距离间隙(Meyer and Paulay, 2005; Meier *et al.*, 2006, 2008),或通过定义一个能区别种间种内关系的阈值,将未知个体鉴定到已知种或定为新种(Blaxter *et al.*, 2005)。基于特征(character-based method)的 DNA 条形码方法,其目的是寻找某分类单元内成员共享、并可以将其与其他分类单元区分开的诊断特征(Neil Sarkar *et al.*, 2002)。一个物种中的固定核苷酸位点可以作为鉴定该种的诊断特征,类似于经典分类学中的形态特征,来鉴定和描绘物种(Rach *et al.*, 2008; Yassin *et al.*, 2010)。这 3 种 DNA 条形码方法均在物种分类的研究中表现出了一定的有效性,但不同方法各有利弊,其在昆虫物种鉴定和识别中同时使用的研究却鲜有报道。

草螟科(Crambidae)属鳞翅目昆虫中的一类小型蛾类,其种类繁多,大量为害各种栽培农作物,往往造成严重损失(王平远,1980)。它们身体细小脆弱,标本采集如不仔细,往往鳞片脱落肢体残缺,给利用传统形态手段的物种鉴定造成一定困难。鉴于此,我们选用北京百花山地区草螟科的物种,利用线粒体 COI 基因的部分序列,进行基于进化树、距离和特征方法的 DNA 条形码分析,不仅为了探索 3 种方法对于草螟科物种识别的可靠性,还可以对 DNA 序列信息在物种鉴定中的应用有一个更加清楚的认知。

1 材料与方法

1.1 标本采集及形态鉴定

本研究中所采用的 88 头草螟科蛾类标本均于 2010 年 7 月 3 (5 日在北京百花山(北纬 39.85°, 东经 115.56°, 海拔 733 m)利用高压汞灯诱集(表 1)。标本采集后低温冷冻致死,展翅保存并制作外生殖器玻片,由草螟科专家陈付强博士进行形态学鉴定。标本保存于首都师范大学生命科学学院遗传多样性与进化课题组。

表 1 样本信息及其COI 序列的 GenBank 登录序列号

Table 1 Crambidae samples used in this study and GenBank accession numbers of their COI gene sequences

属	物种	样本代号	GenBank 登录号
Genus	Species	Sample ID	GenBank accession no.
棘趾野螟属 <i>Anania</i>	元参棘趾野螟 <i>A. verbascalis</i> (Denis & Schiffermüller, 1775)	BHS100703117	KC405045
	元参棘趾野螟 <i>A. verbascalis</i> (Denis & Schiffermüller, 1775)	BHS100703166	KC404978
	元参棘趾野螟 <i>A. verbascalis</i> (Denis & Schiffermüller, 1775)	BHS100704138	KC405046
	元参棘趾野螟 <i>A. verbascalis</i> (Denis & Schiffermüller, 1775)	BHS100704188	KC405001
	元参棘趾野螟 <i>A. verbascalis</i> (Denis & Schiffermüller, 1775)	BHS100704245	KC405044
镰翅野螟属 <i>Circobotys</i>	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100703183	KC405040
	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100703223	KC405043
	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100703243	KC405038
	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100703248	KC405039
	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100704162	KC405041
	横线镰翅野螟 <i>C. heterogenalis</i> (Bremer, 1864)	BHS100705108	KC405042
纵卷叶野螟属 <i>Cnaphalocrocis</i>	稻纵卷叶螟 <i>C. medinalis</i> (Güenée, 1854)	BHS100703207	KC404983
	稻纵卷叶螟 <i>C. medinalis</i> (Güenée, 1854)	BHS100705143	KC405008
草螟属 <i>Crambus</i>	<i>Crambus</i> sp.	BHS100703205	KC404982
	<i>Crambus</i> sp.	BHS100704235	KC405047
中国绢野螟属 <i>Diaphania</i>	四斑绢丝野螟 <i>D. quadrimaculalis</i> (Bremer & Grey, 1853)	BHS100703111	KC405061
	四斑绢丝野螟 <i>D. quadrimaculalis</i> (Bremer & Grey, 1853)	BHS100704179	KC405062
	四斑绢丝野螟 <i>D. quadrimaculalis</i> (Bremer & Grey, 1853)	BHS100704190	KC405063
	四斑绢丝野螟 <i>D. quadrimaculalis</i> (Bremer & Grey, 1853)	BHS100705031	KC405060
斑水螟属 <i>Eoophyla</i>	华斑水螟 <i>E. sinensis</i> (Hampson, 1897)	BHS100704127	KC405011
	华斑水螟 <i>E. sinensis</i> (Hampson, 1897)	BHS100705049	KC405010
切叶野螟属 <i>Herpetogramma</i>	葡萄切叶野螟 <i>H. luctuosalis</i> (Güenée, 1854)	BHS100703126	KC405013
	葡萄切叶野螟 <i>H. luctuosalis</i> (Güenée, 1854)	BHS100703150	KC405016
	葡萄切叶野螟 <i>H. luctuosalis</i> (Güenée, 1854)	BHS100703156	KC405014
	葡萄切叶野螟 <i>H. luctuosalis</i> (Güenée, 1854)	BHS100704113	KC405015
蚀叶野螟属 <i>Lamprosema</i>	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703100	KC404994
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703152	KC404993
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703195	KC404991
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703200	KC404981
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703208	KC404984
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703210	KC404992
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100703219	KC404985
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100704171	KC405004
	黑点蚀叶野螟 <i>L. commixta</i> (Butler, 1879)	BHS100705140	KC405009
	艾锥额野螟 <i>L. aeruginalis</i> (Hübner, 1796)	BHS100704079	KC405031
锥额野螟属 <i>Loxostege</i>	艾锥额野螟 <i>L. aeruginalis</i> (Hübner, 1796)	BHS100704093	KC405029
	艾锥额野螟 <i>L. aeruginalis</i> (Hübner, 1796)	BHS100705119	KC405030
	网锥额野螟 <i>L. sticticalis</i> (Linnaeus, 1761)	BHS100704082	KC404995
	网锥额野螟 <i>L. sticticalis</i> (Linnaeus, 1761)	BHS100704121	KC405032
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)	BHS100703167	KC405035
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)	BHS100703188	KC405034
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)	BHS100703234	KC405033
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)	BHS100704198	KC405036
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)	BHS100705123	KC405037
	枯黄锥额野螟 <i>L. turbidalis</i> (Treitschke, 1829)		

续表 1 (Table 1 continued)

属	物种	样本代号	GenBank 登录号
Genus	Species	Sample ID	GenBank accession no.
扇野螟属 <i>Pleuroptya</i>	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703084	KC405053
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703184	KC405058
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703214	KC405052
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703247	KC404986
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703267	KC405054
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100703269	KC404990
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704086	KC404996
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704118	KC404997
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704143	KC404998
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704148	KC405055
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704149	KC405056
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704155	KC404999
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704193	KC405059
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704197	KC405002
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100704218	KC405051
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705082	KC405049
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705088	KC405048
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705098	KC405050
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705101	KC405005
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705107	KC405006
	三条扇野螟 <i>P. chlorophanta</i> (Butler, 1878)	BHS100705133	KC405057
双突野螟属 <i>Sitochroa</i>	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703085	KC405022
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703087	KC405023
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703089	KC404976
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703103	KC404977
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703139	KC405028
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703174	KC404979
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703176	KC405024
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703189	KC404980
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703258	KC404987
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703259	KC404988
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100703263	KC404989
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100704003	KC405020
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100704010	KC405026
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100704111	KC405019
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100704211	KC405025
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100705084	KC405027
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100705097	KC405017
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100705111	KC405018
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100705118	KC405007
	尖双突野螟 <i>S. verticalis</i> (Linnaeus, 1758)	BHS100705139	KC405021
狭野螟属 <i>Stenia</i>	三环狭野螟 <i>S. charonialis</i> (Walker, 1859)	BHS100704185	KC405000
	三环狭野螟 <i>S. charonialis</i> (Walker, 1859)	BHS100704213	KC405003
	三环狭野螟 <i>S. charonialis</i> (Walker, 1859)	BHS100705121	KC405012

1.2 DNA 提取、PCR 扩增及测序

取样本腿部肌肉组织用 Biomed 组织/细胞基因组 DNA 快速提取试剂盒提取基因组 DNA,然后置于 -20°C 冰箱保存备用。

使用 DNA 条形码通用引物 LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') 和 HCO2198 (5'-TAAACTTCAGGGTGACCAAAAA-ATCA-3') (Folmer *et al.*, 1994) 扩增线粒体 COI 基因 5'端长为 658 bp 的目的片段,采用 25 μL 的 PCR 体系:2 $\times$ Mastermix 12.5 μL ,引物 LCO1490 和 HCO2198 各 0.5 μL (10 $\mu\text{mol/L}$),模板 DNA 2.5 μL , ddH_2O 9 μL 。PCR 反应条件为: 94°C 预变性 5 min;30 个循环包括 94°C 变性 30 s, 50°C 复性 30 s, 72°C 延伸 30 s;最后 72°C 延伸 5 min。PCR 扩增产物经 1% 琼脂糖凝胶电泳检测,由北京中科西林生物公司完成测序。

1.3 数据分析

1.3.1 系统发育分析 使用 Chromas 软件观察序列峰图质量,对每个碱基逐一进行人工校对,经校对的序列用 Clustal X (Thompson *et al.*, 1997) 软件进行序列比对。然后利用 Mega4.0 软件 (Tamura *et al.*, 2007) 基于 Kimura 双参数模型 (Kimura-2-parameter, K2P) (Kimura, 1980) 采用邻接法 (neighbor-joining, NJ) 构建系统发育树,使用 1 000 次重复检测的 Bootstrap 值表示树上各节点的支持率。

1.3.2 距离法分析 利用 Mega4.0 软件基于 K2P 模型计算种内及种间遗传距离,进行“barcoding gap”分析。

使用 ABGD (automatic barcode gap discovery) 软件 (Puillandre *et al.*, 2011) 基于遗传距离对样本进行划分,划分在同一组的样本被认定为是一个种。将 88 个样本的 K2P 遗传距离矩阵在线提交到 ABGD 网站 (<http://www.wabi.snv.jussieu.fr/public/abgd/abgdweb.html>),其中,参数设置均使用缺省值,即种内差异先验值 P (prior intraspecific divergence) 为 0.001 到 0.1,最小相对 gap 宽度值 X (minimum relative gap width) 为 1.5。然后将样本划分结果与其形态鉴定结果比较对照。

1.3.3 特征法分析 本研究基于 14 个草螟物种的 88 个样本,寻找可以作为种级诊断特征的特殊核苷酸位点,并计算成功率。该计算程序由首都

师范大学生命科学院遗传多样性与进化课题组开发 (<http://202.204.209.200/shizishow.php?idh=493>)。

2 结果与分析

2.1 系统发育树聚类结果

以样本的 DNA 为模板,经 PCR 扩增得到 88 条 650 bp 左右的 COI 序列,用 Clustal X 软件比对应后截为 615 bp 的等长片段,序列提交至 GenBank 数据库,检索序列号见表 1。

根据外部形态特征,88 个草螟样本最终被鉴定为 14 个种。使用邻接法以 88 条草螟序列构建 NJ 系统发育树,聚类分析结果如图 1 所示:不同种的草螟被明显的区分,14 种草螟均各自聚成一个单系群,且各单系分支的支持率均为 100%。

2.2 基于距离方法的分类结果

以 K2P 模型计算 14 种草螟 COI 序列的种内与种间遗传距离,分析结果如图 2 所示。种内遗传距离为 0 ~ 1.48%,平均为 0.25%;种间遗传距离为 6.82% ~ 15.69%,平均为 11.38%;每个种的种内最大遗传距离均远远小于种间最小遗传距离,种内和种间存在一个明显的“barcoding gap”区域。

利用 ABGD 软件以 0.001 ~ 0.1 的先验值 P 区间对 88 个草螟样本进行划分,结果包含了初始划分 (initial partition) 和递归划分 (recursive partition) 两种情况,如图 3 所示。其中,初始划分较为稳定,88 个样本均被分成 14 组;递归划分将样本分成了 1 ~ 25 组,但在一个很大的 P 值范围内维持在 14 组。因此我们选择比较稳定的初始划分结果与形态鉴定结果进行比较,发现每个分组与形态种均呈现一一对应的关系 (图 1),基于样本遗传距离的 ABGD 方法鉴定 14 个草螟种的成功率为 100%。

2.3 基于特征方法的条形码分析

通过计算 88 个草螟个体长为 615 bp 的 COI 条形码序列,共找到 74 个诊断核苷酸位点。本研究共 14 种草螟,包括 3 个同属物种,最终在每个种内都找到了各自的诊断核苷酸位点 (表 2),成功率为 100%。这些诊断核苷酸位点在种内比较保守,同时又能将各个种区分开,因此可以根据这些诊断核苷酸位点对草螟物种进行鉴定。本研究

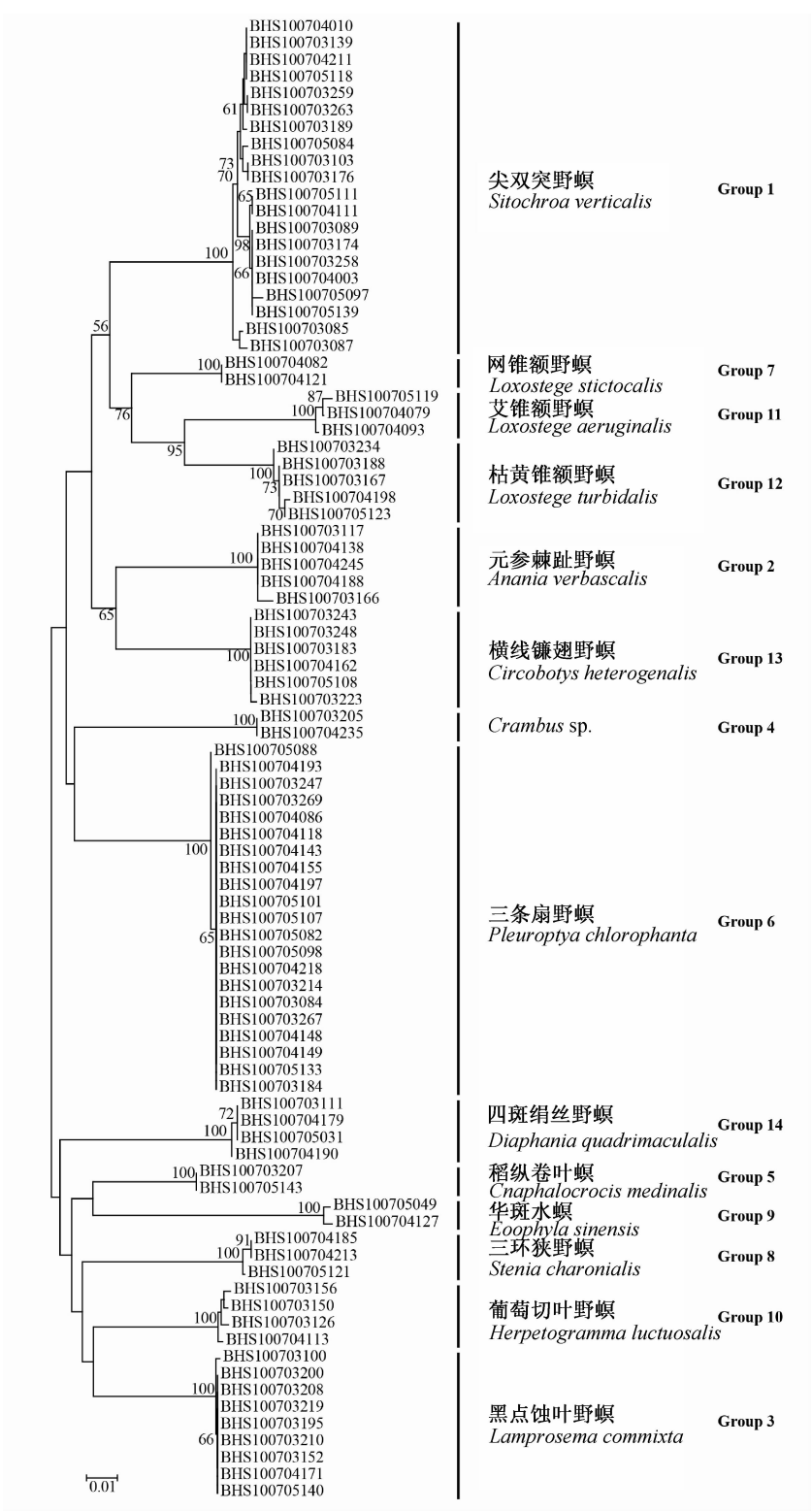


图 1 基于COI 序列构建的百花山地区草螟 NJ 系统发育树

Fig. 1 Neighbour-joining phylogenetic tree of the Crambidae taxa in this study based on COI gene sequences

左侧为草螟样本构建的 NJ 系统发育树,右侧为 ABGD 软件的初始划分结果。

The neighbour-joining phylogenetic tree of the Crambidae taxa is on the left.

The results of initial partition by ABGD soft are on the right.

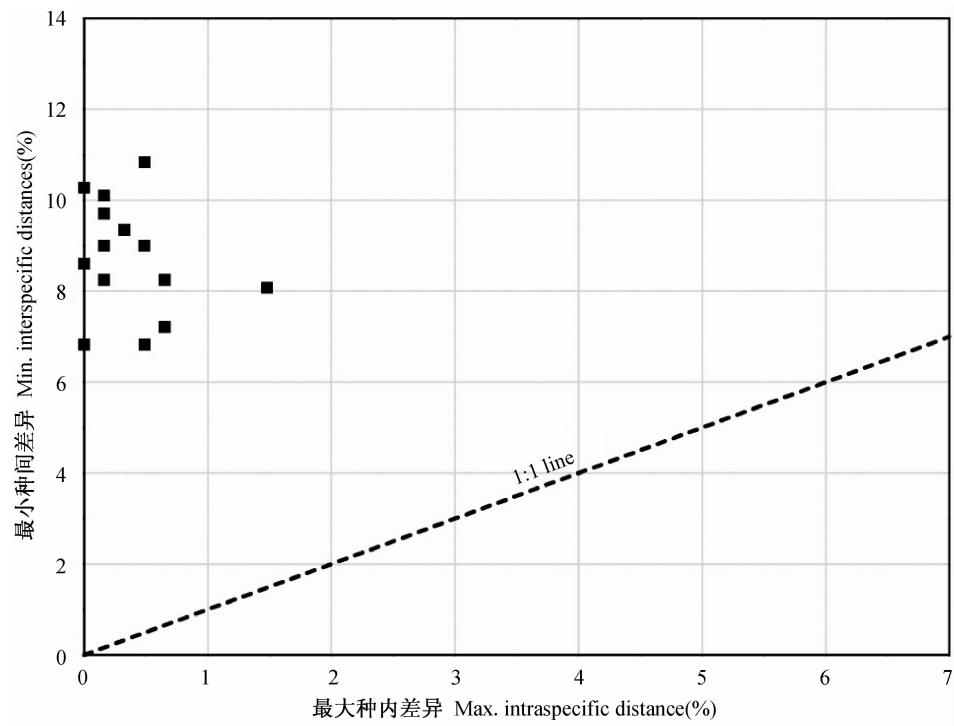


图 2 百花山地区草螟 COI 序列的 DNA barcoding gap 分析

Fig. 2 DNA barcoding gap analysis of COI gene sequence from Crambidae taxa in Baihuashan

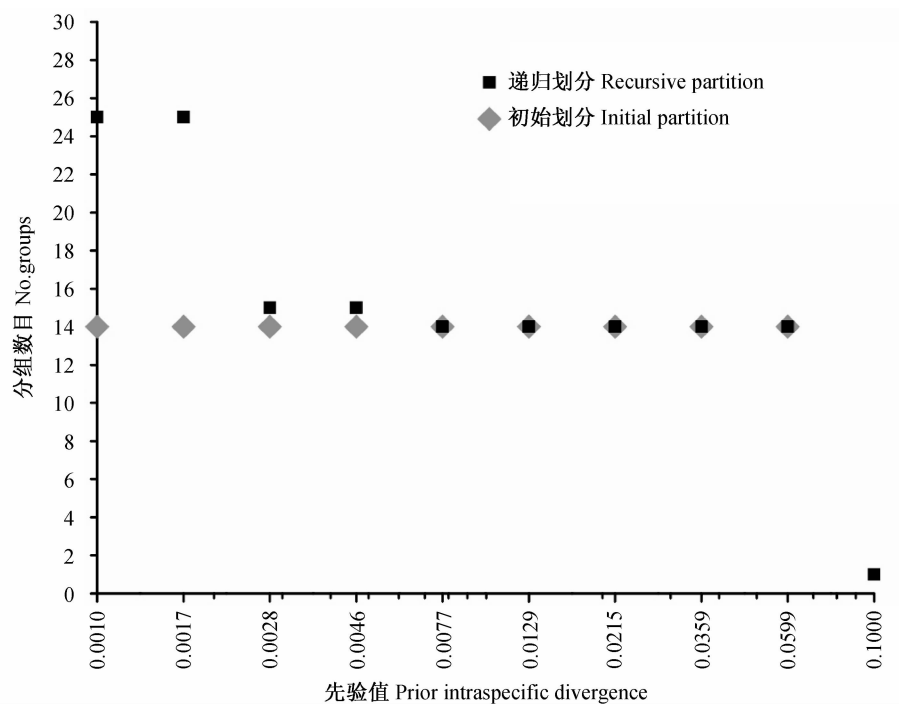


图 3 ABGD 软件对草螟样本的划分结果

Fig. 3 The automatic partition results of Crambidae taxa by ABGD

表 2 14 种草螟基于特征法的 DNA 条形码分析结果夜蛾科样本
Table 2 Character-based DNA barcodes for 14 Crambidae species

物种 Species		诊断特征位点 Diagnostic characters						特征数目 Character numbers	样本数目 Sample numbers
<i>A. verbascalis</i>	C_59 *	C_83	C_197	C_308	T_356	C_512	T_572	7	5
<i>C. heterogenalis</i>	C_2	T_8	C_234	T_263	C_365	C_590		6	6
<i>C. medinalis</i>	T_37	T_188	C_210	T_430				4	2
<i>Crambus</i> . sp.	C_62	G_281	G_440	G_497	C_528	G_578		6	2
<i>D. quadrimaculalis</i>	C_104	C_206	C_350	C_446	C_554			5	4
<i>E. sinensis</i>	C_44	C_92	G_131	G_176	A_423	A_438	T_449	11	2
	G_458	G_530	T_578	C_599					
<i>H. luctuosalis</i>	C_68	C_242	A_296					3	4
<i>L. commixta</i>	C_35	C_215	G_356	C_419				4	9
<i>L. aeruginalis</i>	C_65	A_230	C_323	C_504				4	3
<i>L. sticticalis</i>	T_131	C_173	C_302	C_452				4	2
<i>L. turbidalis</i>	G_107	C_305	C_455	C_569				4	5
<i>P. chlorophanta</i>	C_233	T_269	C_375	C_398	T_575			5	21
<i>Siverticalis</i>	C_152	C_170	C_224	C_383				4	20
<i>S. charonialis</i>	C_149	G_200	T_265	C_316	C_431	C_437	C_497	7	3

注：* C_59:表示第 59 个位点是 C。
* C_59 means the 59th position is C, the same below.

每个种的样本数目不同(2 ~ 21 个), 诊断特征数目也从 3 个到 11 个不等; 含 4 个样本的葡萄切叶野螟 *H. luctuosalis* 诊断核苷酸最少, 含 2 个样本的华斑水螟 *E. sinensis* 诊断核苷酸最多。

3 讨论

DNA 条形码作为一种极为有效的工具在昆虫分类中得了广泛的应用, 它可以及时快速获得分子数据, 极大地简化了物种分类和鉴定工作, 开启了物种分类鉴定的新时代(李青青 等, 2010)。本研究利用基于进化树、距离和特征方法分别对百花山的草螟科蛾类做了条形码分析: 在 88 个草螟样本构建的 NJ 系统发育树上, 14 个种都能各自聚成单系, 不同草螟种类被明显区分; 种间平均遗传距离(11. 38%) 是种内平均遗传距离(0. 25%) 的 45 倍, 完全符合 DNA 条形码有效性的检测标准(Hebert *et al.* , 2003a); 14 个草螟物种均成功地找到了各自的诊断核苷酸位点, 可以作为诊断特征对这些物种进行鉴定。这些结果表明, 3 种方法在草螟科的物种鉴定方面都具有较高的准确性。

在某些类群中, 由于不同物种的线粒体基因碱基置换速率不同, 导致物种种内和种间遗传距

离出现重叠区域(overlap), 从而很难界定区分种内和种间差异的阈值标准(肖金花等, 2004)。本研究使用的 ABGD 方法, 基于一定范围的先验值 P, 自动地探测给定数据集的“barcoding gap”, 在种内和种间有重叠区域时, 依然能够对数据集进行划分。它在一定程度上减少了阈值定义的人为因素, 大大降低了基于距离分类的主观性, 且拥有较高的准确率。我们使用 ABGD 软件对样本分组, 初始划分结果与形态种完全符合; 此外, ABGD 的两种分组方法中, 初始划分较为稳定且与形态分类相似, 这个结果也与 Puillandre 等(2011) 的结论相同。

特征法的原理是在条形码序列中寻找可以区分不同物种的特殊核苷酸位点, 同时忽略同一物种的种内变异, 其在果蝇属(Yassin *et al.* , 2010) 和蜻蜓目(Rach *et al.* , 2008; Damm *et al.* , 2010) 的应用中已取得一定成效。本研究中, 14 个草螟均成功地找到了各自不少于 4 个的诊断特征位点, 表明特征法在草螟科物种鉴定中也有较高的成功率。我们还发现, 样本量较多的物种, 像尖双突野螟 *S. verticalis* (20 头) 和三条扇野螟 *P. chlorophanta* (21 头), 分别只有 4、5 个诊断位点;

而稻纵卷叶螟 *C. medinalis* (2 头)、网锥额野螟 *L. sticticalis* (2 头),也是 4 个诊断位点。从理论上讲,当种内样本量增多时,种内变异也会增大,寻找种内保守位点就会变得比较困难;反之,如果种内样本量较少,跟其余种相比较时,也许找出较多的特异位点数目。本研究中种内样本个数对诊断特征位点的寻找影响不大,也许是因为涉及的物种和样本量较少所致,针对这个问题,尚需收集更多样本进一步做更深入的研究。

基于进化树和距离的 DNA 条形码方法目前广受关注,但也也面临着一些问题和争议 (Moritz and Cicero, 2004; Meyer and Paulay, 2005; Hickerson *et al.*, 2006)。例如不完全谱系分选现象 (incomplete lineage sorting) (Floyd *et al.*, 2002; Wiemers and Fiedler, 2007; McFadden *et al.*, 2011),它导致基因树与物种树发生冲突,构建的基因树不能正确的反映物种间真实的进化关系。这种现象普遍存在于近期分化的物种中 (Van Velzen *et al.*, 2012),由于分化时间较短,这些物种表现出较小的种间遗传距离,在系统发育树上也会形成嵌套关系 (Knowles and Carstens, 2007),基于进化树和距离的条形码方法不能将其准确识别。一些研究表明,基于特征的条形码方法似乎在近缘物种中表现出更高的鉴定效力 (Zou *et al.*, 2011, 2012)。特征法不依赖于遗传距离和拓扑结构,它通过寻找一些种内保守、种间特异的特殊核苷酸诊断位点,体现不同物种之间的差别,从而将近缘物种清晰的区分开来。此外,相对于前两种方法,特征法对于物种的描述更加形象化,可以很好的应用到现在所提倡的“综合分类” (integrative taxonomy) (Dayrat, 2005; Goldstein and DeSalle, 2011)中,即将分子数据和形态生态等传统分类信息相结合进行准确的物种鉴定,从而推进动物分类学以及分子鉴定的快速发展。

本研究基于进化树、距离和特征 3 种方法,进一步充实和拓展了 DNA 条形码在昆虫物种分类方面的研究,结果表明, DNA 条形码能够对草螟科的蛾类实现快速准确的鉴定。但是,本研究中涉及的样本种类较少,每种蛾类的序列数目也有限 (2~21 条),代表性略显不足。我们会继续广泛采集各种鳞翅目标本和收集 DNA 条形码数据,弥补本研究样本量的不足,进一步探讨分析中存在的问题,同时也为鳞翅目条形码数据库的构建贡

献一份力量。

参考文献 (References)

- Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, Abebe E, 2005. Defining operational taxonomic units using DNA barcode data. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360(1462):1935–1943.
- Burns JM, Janzen DH, Hajibabaei M, Hallwachs W, Hebert PDN, 2007. DNA barcodes of closely related (but morphologically and ecologically distinct) species of skipper butterflies (Hesperiidae) can differ by only one to three nucleotides. *J. Lepid. Soc.*, 61(3):138–153.
- Damm S, Schierwater B, Hadrys H, 2010. An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology. *Mol. Ecol.*, 19(18):3881–3893.
- Dayrat B, 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.*, 85(3):407–415.
- Floyd R, Abebe E, Papert A, Blaxter M, 2002. Molecular barcodes for soil nematode identification. *Mol. Ecol.*, 11(4):839–850.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R, 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biotechnol.*, 3(5):294–299.
- Goldstein PZ, DeSalle R, 2011. Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. *Bioessays*, 33(2):135–147.
- Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN, 2006. DNA barcodes distinguish species of tropical Lepidoptera. *PNAS*, 103(4):968–971.
- Hebert PDN, Cywinska A, Ball SL, 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B. Biol. Sci.*, 270(1512):313–321.
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W, 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *PNAS*, 101(41):14812–14817.
- Hebert PDN, Ratnasingham S, DeWaard JR, 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B. Biol. Sci.*, 270(Suppl. 1):S96–S99.
- Hickerson MJ, Meyer CP, Moritz C, 2006. DNA barcoding will often fail to discover new animal species over broad parameter space. *Syst. Biol.*, 55(5):729–739.
- Kimura M, 1980. A simple method for estimating evolutionary

- rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 16(2):111–120.
- Knowles LL, Carstens BC, 2007. Delimiting species without monophyletic gene trees. *Syst. Biol.*, 56(6):887–895.
- McFadden CS, Benayahu Y, Pante E, Thoma JN, Nevarez PA, France SC, 2011. Limitations of mitochondrial gene barcoding in Octocorallia. *Mol. Ecol. Resour.*, 11(1):19–31.
- Meier R, Shiyang K, Vaidya G, Ng PKL, 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.*, 55(5):715–728.
- Meier R, Zhang G, Ali F, 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Syst. Biol.*, 57(5):809–813.
- Meyer CP, Paulay G, 2005. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.*, 3(12):e422.
- Moritz C, Cicero C, 2004. DNA barcoding: promise and pitfalls. *PLoS Biology*, 2(10):e354.
- Neil Sarkar I, Thornton JW, Planet PJ, Figurski DH, Schierwater B, DeSalle R, 2002. An automated phylogenetic key for classifying homeoboxes. *Mol. Phylogenet. Evol.*, 24(3):388–399.
- Puillandre N, Lambert A, Brouillet S, Achaz G, 2011. ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.*, 21(8):1864–1877.
- Rach J, DeSalle R, Sarkar IN, Schierwater B, Hadrys H, 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. *Proc. R. Soc. Lond. B. Biol. Sci.*, 275(1632):237–247.
- Radulovici AE, Marie BS, Dufresne F, 2009. DNA barcoding of marine crustaceans from the Estuary and Gulf of St Lawrence: a regional-scale approach. *Mol. Ecol. Resour.*, 9(S1):181–187.
- Saunders, GW, 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360(1462):1879–1888.
- Smith MA, Fisher BL, Hebert PDN, 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360(1462):1825–1834.
- Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, 24(8):1596–1599.
- Tavares ES, Baker AJ, 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evol. Biol.*, 8(1):81.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, 25(24):4876–4882.
- Van Velzen R, Weitschek E, Felici G, Bakker FT, 2012. DNA barcoding of recently diverged species: relative performance of matching methods. *PLoS ONE*, 7(1):e30490.
- Ward RD, Hanner R, Hebert PDN, 2009. The campaign to DNA barcode all fishes, FISH-BOL. *J. Fish Biol.*, 74(2):329–356.
- Wiemers M, Fiedler K, 2007. Does the DNA barcoding gap exist? -a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front. Zool.*, 4(8):16.
- Wong EHK, Hanner RH, 2008. DNA barcoding detects market substitution in North American seafood. *Food Res. Int.*, 41(8):828–837.
- Yassin A, Markow TA, Narechania A, O’Grady PM, DeSalle R, 2010. The genus *Drosophila* as a model for testing tree- and character-based methods of species identification using DNA barcoding. *Mol. Phylogenet. Evol.*, 57(2):509–517.
- Zou SM, Li Q, Kong LF, 2012. Monophyly, distance and character-based multigene barcoding reveal extraordinary cryptic diversity in nassarius: A complex and dangerous community. *PLoS ONE*, 7(10):e47276.
- Zou SM, Li Q, Kong LF, Yu H, Zheng XD, 2011. Comparing the usefulness of distance, monophyly and character-based DNA barcoding methods in species identification: a case study of Neogastropoda. *PLoS ONE*, 6(10):e26619.
- 李青青, 李地艳, 段焰青, 李加敏, 刘晓飞, 曹能, 叶辉, 2010. DNA 条形码在鳞翅目昆虫中的应用. *生命科学*, 22(4):307–312.
- 马英, 鲁亮, 2010. DNA 条形码技术研究新进展. *中国媒介生物学及控制杂志*, 21(3):275–280.
- 马英, 鲁亮, 2012. DNA 条形码技术. *中国媒介生物学及控制杂志*, 23(3):185–190.
- 王平远, 1980. 中国经济昆虫志(鳞翅目: 螟蛾科). 北京: 科学出版社. 1–229.
- 肖金花, 肖晖, 黄大卫, 2004. 生物分类学的新动向—DNA 条形码 3. *动物学报*, 50(5):852–855.