

切叶蚁亚科七属十二种的分子系统学研究

陈振鹏, 周善义*

(广西师范大学生命科学学院 广西桂林 541004)

摘要: 测定了切叶蚁亚科 7 属 12 种的线粒体 CO1、CO2 的部分基因及完整的 tRNA_{Leu} 基因 DNA 序列, 对 DNA 序列进行了分析, 对 tRNA_{Leu} 基因进行了二级结构分析; 根据 DNA 序列数据和氨基酸序列数据, 以臭蚁亚科的 *Forelius chalybaeus* 作为外群, 采用最大似然法 (ML)、最大简约法 (MP)、邻接法 (NJ)、未加权组对算术平均法 (UPGMA) 构建分子系统树, 通过自举检验, 得到自举置信水平, 以此检验该分子系统树的可靠性。研究结果显示, 基于以上基因的分子系统分析与传统分类分析的结果基本一致, 且在属级的一致性高于种级的一致性。

关键词: 切叶蚁亚科; 线粒体 DNA; 基因; 系统发育

中图分类号: Q968 文献标识码: A 文章编号: 0454-6296(2007)04-0395-10

Molecular systematic study on twelve species of seven genera in Myrmicinae (Hymenoptera: Formicidae) from Guangxi, South China

CHEN Zhen-Peng, ZHOU Shan-Yi* (College of Life Science, Guangxi Normal University, Guilin, Guangxi 541004, China)

Abstract: Cytochrome oxidase subunits 1 (CO1) and 2 (CO2), and the entire sequence of the gene for tRNA leucine from twelve species of seven genera of the ant subfamily Myrmicinae were sequenced. The DNA sequences were analyzed, and the secondary structure of tRNA leucine also gene was studied. The molecular phylogenetic analyses of the subfamily Myrmicinae with *Forelius chalybaeus*, a species of the subfamily Dolichoderinae as the outgroup were performed by using four methods: the maximum likelihood (ML), maximum parsimony (MP), neighbor-joined (NJ), and unweighted pairgroup method with arithmetic means (UPGMA) based on the deoxyribonucleic acid data and the amino acid data. The taxonomic scheme for these genera and species based on molecular phylogeny of above genes consisted with that of the traditional taxonomy and the consistency for genera was better than that for species.

Key words: Subfamily Myrmicinae; mtDNA; gene; phylogeny

蚂蚁隶属于膜翅目蚁科 (Hymenoptera: Formicidae), 是典型的社会性昆虫。据估计, 全世界的蚂蚁有 15 000 种, 形成庞大的动物群体, 在陆地生态系统中占有优势地位 (Wilson, 1992)。蚂蚁体现出高度的社会行为, 使其在进化生物学研究中显得相当重要 (Hölldobler and Wilson, 1990; Bourke and Franks, 1995; Corizer and Pamilo, 1996)。国外学者 Crozier 等 (1993, 1995, 1996, 1997) 对蚂蚁进行过大量的分子系统发育研究, 根据线粒体 DNA 序列进行亚科级水平上的分子系统发育研究, 还用细胞色素 b DNA 序列分析了澳大利亚跳蚁 *Myrmecia pilosula*

种组之间的亲缘关系, 所得结果与核型研究的结果十分吻合。

到目前为止我国尚未有过蚂蚁分子系统学研究的报道。切叶蚁亚科是蚁科中最大的亚科, 通过分子手段从微观上探索其系统发育关系, 找到适合于不同分类阶元的分子生物学方法和合适的基因片段, 可以为研究我国切叶蚁亚科昆虫的分子系统学研究积累资料。本研究以 7 属 12 种切叶蚁亚科昆虫为研究对象, 应用 DNA 序列测定技术获得样本的线粒体细胞色素氧化酶亚基 1 和 2 及完整的 tRNA_{Leu} 基因 DNA 序列; 运用 MEGA、DNASP、Phylip

基金项目: 国家自然科学基金项目 (30360013); 广西科学基金项目 (0236015)

作者简介: 陈振鹏, 男, 1979 年生, 硕士, 研究方向为昆虫分子生物学

* 通讯作者 Author for correspondence, E-mail: syzhou@mailbox.gxnu.edu.cn

收稿日期 Received: 2006-06-12; 接受日期 Accepted: 2007-03-01

软件对这些基因序列数据进行处理分析,以臭蚁亚科的 *Forelius chalybaeus* 作为外群,构建系统树,从分子水平上探讨它们之间的分类地位和亲缘关系。

1 材料与方法

1.1 样本来源

本研究所用样本采自广西壮族自治区,共 7 属 12 种,所有标本均保存于无水乙醇中,置于 -70°C 条件下保存(表 1)。

1.2 基因组 DNA 的提取与检测

挑选保存完整的无水乙醇浸制标本,去其腹部后迅速研磨粉碎,匀浆液浸泡于水浴中,然后按传统的酚-氯仿法提取总 DNA,溶解后分装,于 -20°C 保存备用。对于所提取的 DNA 样品用凝胶电泳检测法和紫外分光光度计检测法进行定量检测。

1.3 目的基因扩增

1.3.1 PCR 扩增引物:PCR 扩增引物参照 Chiotis 等(2000)所用引物(表 2)。

表 1 研究样本

Table 1 Specimens used and their provenances

属 Genera	种名 Species	采样地点及时间 Locality and collecting date	CO1, CO2, rRNA _{Leu} 基因 GenBank 登录号 GenBank accession no. of CO1, CO2 and tRNA _{Leu} genes
大头蚁属 <i>Pheidole</i>	卡氏大头蚁 <i>P. capellinii</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762645
	史氏大头蚁 <i>P. Smythiesii</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762651
	宽结大头蚁 <i>P. noda</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762650
盘腹蚁属 <i>Aphaenogaster</i>	史氏盘腹蚁 <i>A. Smythiesii</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762646
	盘腹蚁 <i>A. sp.</i>	广西环江县, 2-VIII-2003 Huanjiang County, Guangxi, 2 August, 2003	AY762655
红蚁属 <i>Myrmica</i>	中华红蚁 <i>M. sinensis</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762652
	角红蚁 <i>M. angulata</i>	广西罗城县, 26-VII-2003 Luocheng County, Guangxi, 26 July, 2003	AY762656
铺道蚁属 <i>Tetramorium</i>	铺道蚁 <i>T. caespitum</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762647
	双脊铺道蚁 <i>T. bicarinatum</i>	广西罗城县, 26-VII-2003 Luocheng County, Guangxi, 26 July, 2003	AY762653
棱胸切叶蚁属 <i>Pristomyrmex</i>	棱胸双针蚁 <i>P. pungens</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762649
举腹蚁属 <i>Crematogaster</i>	游举腹蚁 <i>C. vagula</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762648
巨首蚁属 <i>Pheidologeton</i>	黑沟巨首蚁 <i>P. melasolenus</i>	广西兴安县, 12-VII-2003 Xing'an County, Guangxi, 12 July, 2003	AY762657

表 2 PCR 所用引物

Table 2 Primers used in PCR

J2791	ATACCTCGACGTTATTCAGA	3093 - 3112 →
H3665	CCACAAATTCTGAACATTG	← 4194 - 4213
C2-N-3661 (Barbara)	CCACAAATTCTGAACATTGACCA	← 4190 - 4213

注:根据实验情况,引物 C2-N-3661(Barbara)可被引物 H3665 所代替。

Notes: The primer C2-N-3661(Barbara) may be replaced by the primer H3665.

1.3.2 PCR 反应体系:每一样品的 PCR 反应总体积为 25 μL , 20 ~ 50 ng 模板 DNA, 2 U Taq 酶, $1 \times$ PCR Buffer, 2.0 mmol/L MgCl_2 , 0.2 mmol/L dNTPs, 0.5 $\mu\text{mmol/L}$ PCR 引物。PCR 反应条件:运行 34 个循环,每一个循环包括 94°C 变性 1 min, 根据实际情

和不同物种,退火温度在 40°C ~ 52°C 范围内调整,退火时间在 30 s ~ 1 min 45 s 范围内调整, 72°C 延伸 2 min, 在正式循环前进行 94°C 预变性 5 min, 循环完毕后 72°C 延伸 5 min。PCR 产物用 1.5% 的琼脂糖凝胶电泳进行检测。

1.4 PCR 产物回收、测序

对于扩增效果良好且足量的样品进行回收。回收产物浓度达到 100 ng/ μ L 的样品委托大连宝生物有限公司进行 DNA 测序。序列分析仪为 ABI 公司 377 型全自动序列分析仪。

1.5 数据处理

使用 ClustalX 软件对上述测定的 DNA 序列进行连配,然后向 NCBI 提交 DNA 序列,登录号如表 1 所示。通过 tRNAscan-SE 来确定 tRNA_{Leu} 序列,用分子进化遗传分析软件 MEGA 和 PHYLIP 软件包,用已发表的 *Forelius chalybaeus*(Chiotis *et al.*, 2000) 的线粒体基因(mtDNA)中相关的 DNA 序列(GenBank accession no. AF147047)作为分类外群,采用最大似然法(ML)、最大简约法(MP)、邻接法(NJ)、未加权组对算术平均法(UPGMA)构建分子系统树,并通过自

举检验,得到自举置信水平,以此来检验该分子系统树的可靠性。

2 结果与分析

2.1 DNA 序列组成

通过对研究所得的 DNA 序列进行碱基成分分析,CO1 基因部分片段:G + C 占 25.23%,A + T 占 74.77%;CO2 基因部分片段:G + C 占 24.56%,A + T 占 75.44%,整个 DNA 序列 G + C 占 24.25%,A + T 占 75.75%,总体上明显表现出 AT 偏倚特征,这与其他昆虫类似。

2.2 DNA 片段中包含的 IGS(intergenic spacer)区

研究中有些物种存在 IGS 区,并且 IGS 区位于 CO1 和 tRNA_{Leu} 之间(表 3)。

表 3 DNA 片段中包含的 IGS(intergenic spacer)区

Table 3 Intergenic spacer region included in DNA sequence

属 Genera	种名 Species	序列 Sequence	长度 Length (bp)
大头蚁属 <i>Pheidole</i>	卡氏大头蚁 <i>P. capellinii</i>	AATAATTTGTATAAAAATTTACTTAA	25
	宽结大头蚁 <i>P. noda</i>	TTTATTTTATATTATATTACTTAA	25
	史氏大头蚁 <i>P. smythiesii</i>	TTTTTTTTAACTTATATATTTAA	22
隆胸切叶蚁属 <i>Pristomyrmex</i>	双针蚁 <i>P. pungens</i>	TTAATATTATTTATAATTACATAA	23
	红蚁属 <i>Myrmica</i>		
红蚁属 <i>Myrmica</i>	中华红蚁 <i>M. sinensis</i>	AGATTTTAA	9
	角红蚁 <i>M. angulata</i>	AGATTTTAA	9
盘腹蚁属 <i>Aphaenogaste</i>	史氏盘腹蚁 <i>A. smythiesii</i>	AT	2
	盘腹蚁 <i>A. sp.</i>	AT	2

2.3 DNA 序列多样性分析

同义(或沉默替代):CO2 基因的同义或沉默替代其位点数目为 107.64,大于 CO1 基因的同义或沉默替代其位点数目为 33.67;CO2 基因的 Pi(s)(0.39346,0.58066)大于 CO1 基因的 Pi(s)(0.08478,0.09156),CO2 基因的 Theta(s)0.30764 大于 CO1 基因的 Theta(s)0.04918;CO2 基因发生的突变数目(number of mutations)100,远大于 CO1 基因发生的突变数目 5。

非同义(或氨基酸更换替代):CO2 基因的同义或沉默替代其位点数目为 444.36,大于 CO1 基因的同义或沉默替代其位点数目为 140.33;CO2 基因的 Pi(s)(0.13773,0.15362)大于 CO1 基因的 Pi(s)(0.23777,0.29184),CO2 基因的 Theta(s)0.05962 大于 CO1 基因的 Theta(s)0.09439;CO2 基因发生的突变数目 80,大于 CO1 基因发生的突变数目 40。

2.4 密码子使用频率[CU]和相对使用频率[RSCU]分析

富含 AT 的密码子:Phe(F)的 CF 值和 RSCU 值范围为 11.7(0.86)~15.3(1.14);Leu(L)的 CF 值

和 RSCU 值范围为 0.7(0.10)~18.3(2.76);Ile(I)的 CF 值和 RSCU 值范围为 6.7(0.60)~15.4(1.40);Met(M)的 CF 值和 RSCU 值范围为 7.3(0.77)~11.6(1.23);Tyr(Y)的 CF 值和 RSCU 值为 2.4(0.32)~12.5(1.68);Asn(N)的 CF 值和 RSCU 值范围为 2.7(0.33)~13.6(1.67);Lys(K)的 CF 值和 RSCU 值范围为 2.8(0.40)~11.0(1.60)。

富含 GC 的密码子 Pro(P)的 CF 值和 RSCU 值范围为 0.4(0.16)~4.1(1.53);Ala(A)CF 值和 RSCU 值范围为 0.0(0.00)~1.1(2.08);Gly(G)的 CF 值和 RSCU 值范围为 0.2(0.24)~1.1(1.53);Arg(R)的 CF 值和 RSCU 值范围为 0.2(0.17)~2.4(2.47)。

通过对 CO1 基因和 CO2 基因密码子使用频率[CU]和相对使用频率[RSCU]分析,其结果显示富含 AT 的密码子的使用频率(包括相对使用频率)都远大于富含 GC 的密码子使用频率(包括相对使用频率)。

2.5 种群间转换(transition)与颠换(transversion)分析

基因突变中的转换(transition)或颠换

(transversion) 在属内,即种间差异很小,甚至是没有差异,如史氏盘腹蚁 *A. smythiesii* 对盘腹蚁 *A. sp.* 的转换或颠换值均为 0.000,而转换或颠换值在属间差异比较明显。黑沟巨首蚁 *P. melasolenus* 对举脊腹蚁 *C. vagula* 转换偏相关距离值为 0.161,黑沟巨首蚁 *P. melasolenus* 对中华红蚁 *M. sinensis* 颠换偏相关距离值为 0.160。而且从整体上来看,转换大于颠换,其偏相关距离平均值为 1.076。这个特征与其他文献报道相一致(Fitch and Margoliash, 1967; Gojoberi *et al.*, 1982; Kocher and Wilson, 1991) 并且 CO1 基因的转换和颠换偏相关距离平均值为 0.125 和 0.138,均高于 CO2 基因的转换和颠换偏相关距离平均值为 0.105 和

0.119,并且 CO1 基因的 d 值(0.263)高于 CO2 基因的 d 值(0.224),但 CO1 基因的 R 值(0.929)却低于 CO2 基因的 R 值(1.040)。

2.6 tRNA_{Leu}基因二级结构分析

对 tRNA_{Leu} 基因各个部分的二级结构特征的定义分析,用于检测拓扑结构的合理性。tRNA_{Leu} 基因二级结构主要包括四臂四环,它们分别是:DHU (DHU arm)臂与 DHU 环(DHU loop);TΨC 臂(TΨC arm)与 TΨC 环(TΨC loop);河变环(Variable loop);氨基酸臂(其 3' 端为接受氨基酸的部位,即 CpCpAOH);反密码子环(Anticodon loop)和臂。本研究物种的 tRNA_{Leu} 基因二级结构如图 1 所示。

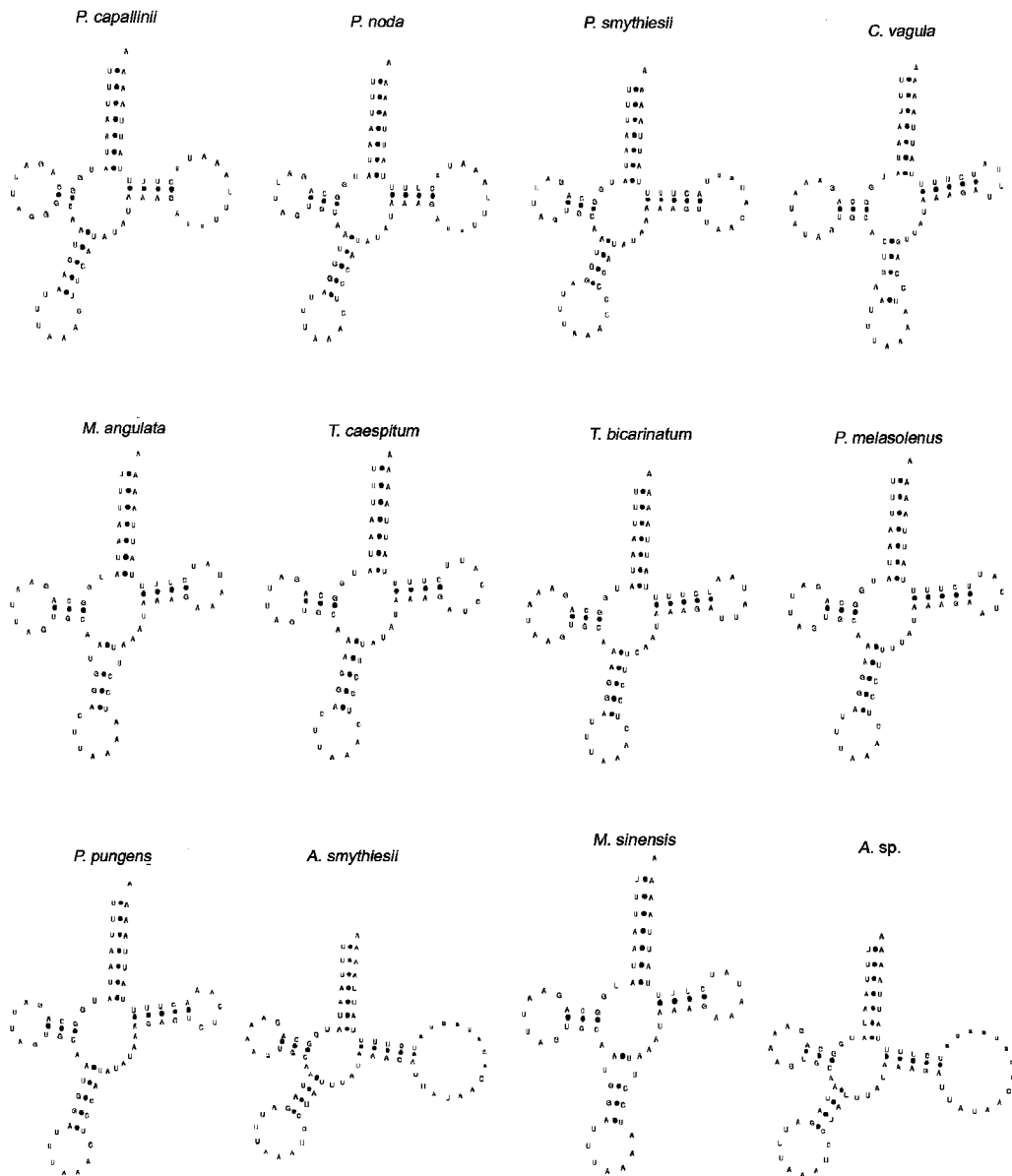


图 1 tRNA_{Leu} 基因二级结构

Fig. 1 The secondary structure of the tRNA_{Leu} gene

在所研究的物种中,CO2 的起始密码子大部分为 ATT,游举腹蚁 *C. vagula* 起始密码子为 ATC,通常这两种密码子都产生异亮氨酸;铺道蚁 *T. caespitum* 起始为甲硫氨酸标准的密码子 ATA;对本研究所用物种的 CO1 基因的终止密码子分析结果显示,铺道蚁属的 *T. caespitum* 和 *T. bicarinatum*、举腹蚁属的 *C. vagula*、红蚁属的 *M. angulata*、巨首蚁属的 *P. melasolenus* 终止密码子为 T,需要通过 mRNA 的 3' 端上冗余的 A 碱基额外补充才能完成

完整的终止密码子;其他的物种则都有完整的终止密码 TA。

2.7 用 CO1、CO2 和 tRNA_{Leu} 3 个基因的 DNA 数据和氨基酸数据构建并检验分子系统树

基于 CO1、CO2 和 tRNA_{Leu} 3 个基因的 DNA 数据和氨基酸数据,用 ML, MP, NJ, UPGMA 方法构建的分子系统树均采用 Tamura-Nei 模型,并做自举检验检测分子系统树的可靠性,结果如图 2 和图 3 所示。

ML 树结果表明:中华红蚁 *M. sinensis* + 角红

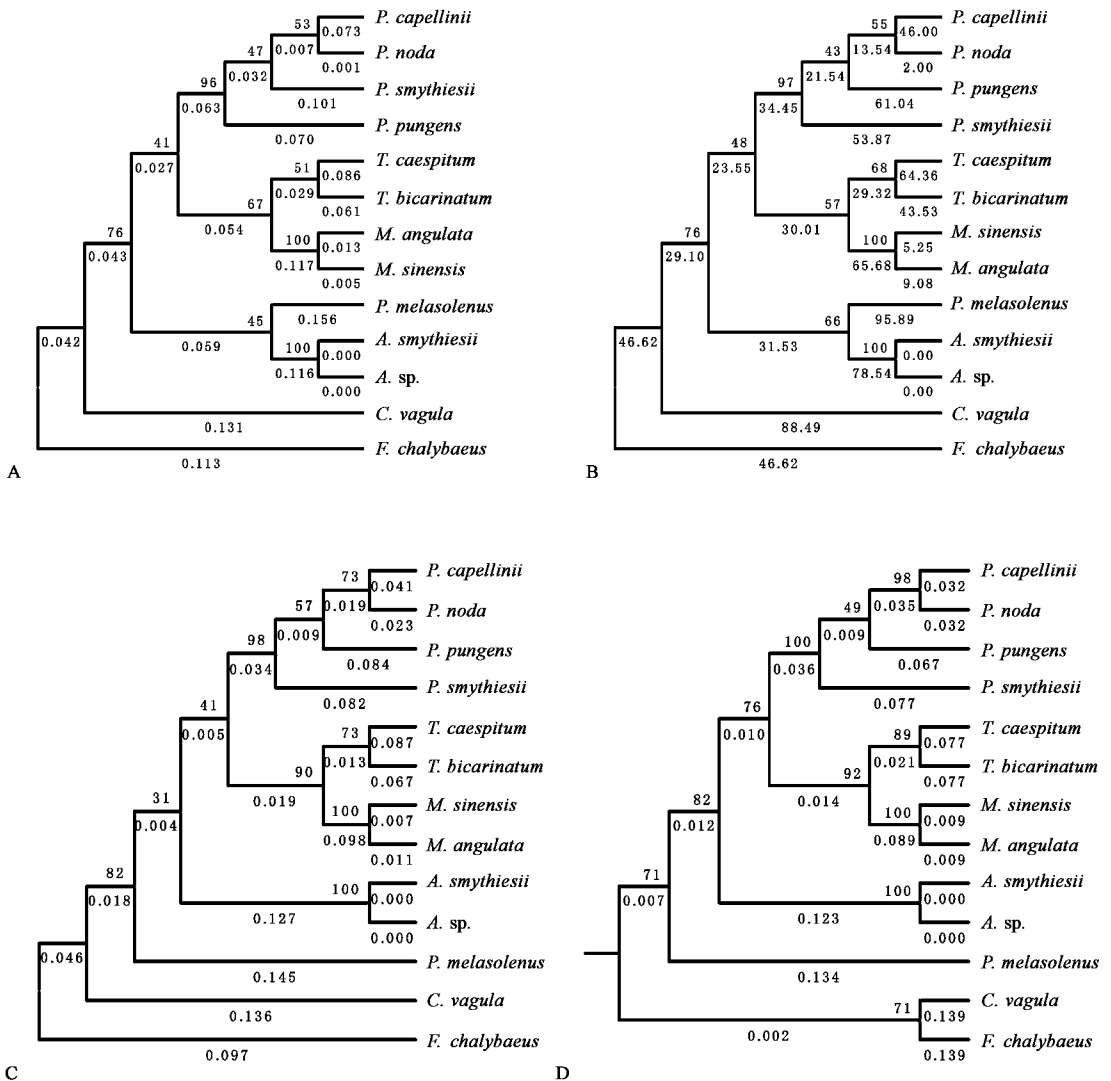


图 2 基于 DNA 数据用 4 种方法构建的切叶蚁亚科 7 属 12 种的分子系统树
(以臭蚁亚科的 *Forelius chalybaeus* 作为外群)

Fig. 2 Phylogenetic trees of 12 species of 7 genera in the subfamily Myrmicinae based on DNA data set with *Forelius chalybaeus*, a species of the subfamily Dolichoderinae as the outgroup

A: 最大似然法树 Maximum likelihood tree; B: 最大简约法树 Maximum parsimony tree; C: 邻接法树 Neighbor-joined tree; D: 未加权组对算术平均法树 UPGMA tree. 下同。The same for Fig. 3.

蚁 *M. angulata* 的 BCL(自举置信水平 boot confidence level, 简称 BCL) 值为 100; 铺道蚁 *T. caespitum* + 双脊铺道蚁 *T. bicarinatum* 的 BCL 值为 51; 红蚁属 + 铺道蚁属 BCL 值为 67; 卡氏大头蚁 *P. capellinii* + 宽结大头蚁 *P. noda* 的 BCL 值为 53; 史氏大头蚁 *P. smythiesii* + 卡氏大头蚁 *P. capellinii* + 宽结大头蚁 *P. noda* 的 BCL 值只有 47; 棱胸切叶蚁属的双针蚁 *P. pungens* + 大头蚁属 3 个物种 BCL 值为 96; 红蚁属 + 铺道蚁属 + 大头蚁属 + 棱胸切叶蚁属的双针蚁 *P. pungens* 的 BCL 值为 41; 氏盘腹蚁 *A. smythiesii* + 盘腹蚁 *A. sp.* 的 BCL 值为 100; 巨首蚁属的黑沟巨首蚁 *P. melasolenus* + 盘腹蚁属 BCL 值为 45; 举腹蚁属的游举腹蚁 *C. vagula* + 其他物

种的 BCL 值为 76; NJ 树结果与 ML 树的结果相比较, 其差异主要是: 史氏大头蚁 *P. smythiesii* 与棱胸切叶蚁属的双针蚁 *P. pungens* 在分支上的位置发生了互换, 并引起了卡氏大头蚁 *P. capellinii* + 宽结大头蚁 *P. noda* 的 BCL 值发生变化, 由 53 升高到 93; MP 树与 ML 树区别不大。UPMGA 树与 ML 树差异主要是: 棱胸切叶蚁属的双针蚁 *P. pungens* 与卡氏大头蚁 *P. capellinii* 的位置发生互换, 举腹蚁属的游举腹蚁 *C. vagula* 取代了巨首蚁属的黑沟巨首蚁 *P. melasolenus* 在 ML 树分支中的位置, 但其 BCL 值仅有 37, 而巨首蚁属的黑沟巨首蚁 *P. melasolenus* 与红蚁属 + 棱胸切叶蚁属 + 大头蚁属组成一个大分支, 但该分支的 BCL 值只有 28。

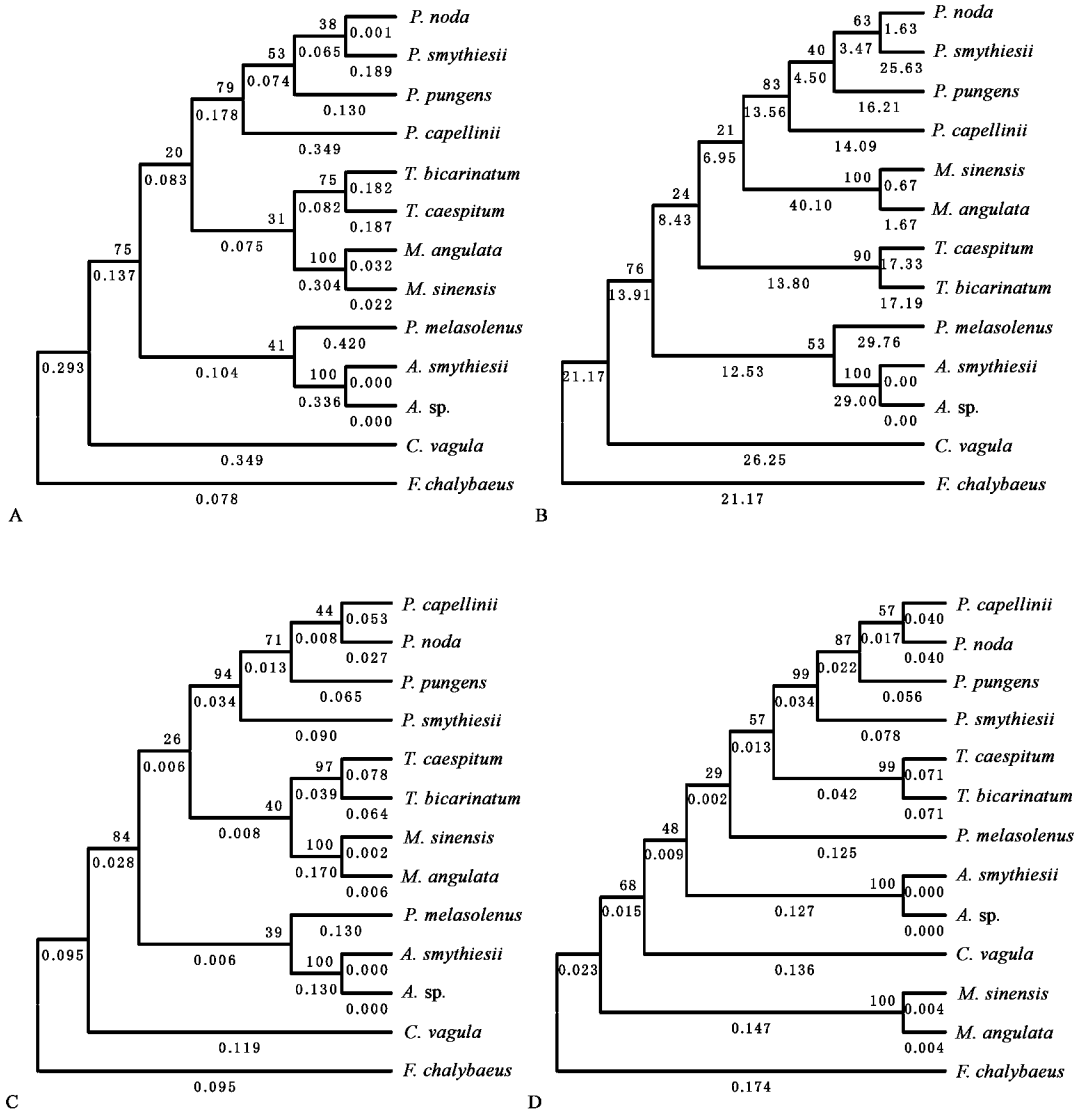


图 3 基于氨基酸数据通过 4 种方法构建的切叶蚁亚科 7 属 12 种的分子系统树

Fig. 3 Phylogenetic trees of 12 species of 7 genera in the subfamily Myrmicinae based on the amino acid data

ML 树显示,史氏大头蚁 *P. smythiesii* + 宽结大头蚁 *P. noda* 的 BCL 值为 37;宽结大头蚁 *P. noda* + 史氏大头蚁 *P. smythiesii* + 棱胸切叶蚁属的双针蚁 *P. pungens* 的 BCL 值只有 53;卡氏大头蚁 *P. capellinii* + 棱胸切叶蚁属的双针蚁 *P. pungens* + 大头蚁属的另外 2 个物种的 BCL 值为 77;双脊铺道蚁 *T. bicarinatum* + 铺道蚁 *T. caespitum* 的 BCL 值为 75;中华红蚁 *M. sinensis* + 角红蚁 *M. angulata* 的 BCL 值为 100;红蚁属 + 铺道蚁属 BCL 值只有 32;大头蚁属 + 棱胸切叶蚁属的双针蚁 *P. pungens* + 红蚁属 + 铺道蚁属 BCL 值仅为 20;史氏盘腹蚁 *A. smythiesii* + 盘腹蚁 *A. sp.* 的 BCL 值为 100;巨首蚁属的黑沟巨首蚁 *P. melasolenus* + 盘腹蚁属的 BCL 值只有 41;举腹蚁属的游举腹蚁 *C. vagula* + 其他物种 BCL 值有 75。NJ 树与 ML 树的区别在于:史氏大头蚁 *P. smythiesii* 与卡氏大头蚁 *P. capellinii* 的位置发生了互换,而且其 BCL 值由 77 升高到 90。MP 树与 ML 树结果基本一致,只是有些分支的 BCL 值发生了变化。UPGMA 树与 NJ 树较相似,但史氏大头蚁 *P. smythiesii* 与棱胸切叶蚁属的双针蚁 *P. pungens* 位置互换,且使得这些分支的 BCL 值提高到 99,97。另外,巨首蚁属的黑沟巨首蚁 *P. melasolenus* 不再与盘腹蚁属组成一个小分支,而是整个大分支的较外侧。举腹蚁属的游举腹蚁 *C. vagula* 在整个系统树最外侧(分类外群 *F. chalybaeus* 除外)。

3 讨论

本研究显示,CO2 基因的 A + T 含量高于 CO1 基因的 A + T 含量。由于 A + T-rich 区主要位于包含该基因复制起点的控制区,而对于 tRNA 来说,A + T-rich 区变异程度最高,即 DNA 片段包含的 A + T-rich 区可能越多,该 DNA 片段进化也就越快(Clary and Wolstenholme, 1985; Fauron and Wolstenholme, 1990; Pashley and Ke, 1992)。传统分类认为,大头蚁属和红蚁属是切叶蚁亚科中较进化的类群,而举腹蚁属、巨首蚁属和盘腹蚁属是较古老的类群。本研究涉及的切叶蚁亚科各个属之间的 A + T 含量存在差异,其中大头蚁属的 A + T 含量最高,卡氏大头蚁、宽结大头蚁和史氏大头蚁的 A + T 含量分别为 87.5%、78.6% 和 75.0%,平均含量达 80%;史氏大头蚁过去归属于四节大头蚁属 *Ceratopheidole*,进化地位较大头蚁属低,故其 A + T 含量较低;中华红蚁

和角红蚁外形上十分相似,其 A + T 含量也非常接近,分别为 78.9% 和 78.5%;史氏盘腹蚁和盘腹蚁的未定名种的外形十分相近,其 A + T 含量也非常接近,分别为 74.7% 和 74.9%;较原始的类群黑沟巨首蚁和游举腹蚁的 A + T 含量最低,分别为 71.8% 和 71.7%。A + T 含量反映出的进化关系与传统分类学的结论基本一致。根据以上所述观点推测,7 个属在系统发育上由原始到进化的顺序为:举腹蚁属 ≥ 巨首蚁属 > 棱胸切叶蚁属 > 盘腹蚁属 > 铺道蚁属 > 大头蚁属 > 红蚁属。通过 CO1 与 CO2 基因 A + T 含量对比,根据上述观点推测 CO2 基因的进化速率大于 CO1 基因。

IGS (intergenic spacer) 区属于 NCS (noncoding sequences) 中的一种,在某种程度上,NCS 也是生物从非细胞 → 单细胞 → 多细胞进化的重要尺表。在真核生物的基因组中,基因只占很少的比例,大部分甚至绝大部分是 NCS,这似乎意味着非编码序列的出现代表了基因组在某种程度上的进化(齐义鹏, 1998)。本研究中有物种存在 IGS 区,位于 CO1 和 tRNA_{Lys} 之间,据此推测本研究物种由原始到进化的排列顺序大致为:举腹蚁属 ≥ 巨首蚁属 ≥ 铺道蚁属 > 盘腹蚁属 > 红蚁属 > 棱胸切叶蚁属 > 大头蚁属。

用两种方法分析所得的结果,铺道蚁属和棱胸切叶蚁属的进化地位差异较大,其原因还有待于今后进一步研究。

由于密码子的简并性 (degeneracy) 和摆动性 (wobble),大多数同义替代发生在密码子的第三位上,而只有少量出现在第一位上。在第二位的所有替代不是同义替代就是无义替代。假定基因组中所有密码子以同一频率出现,且所有核苷酸的替代概率相同,则同义、非同义和无义替代的百分比大致为 25%、71% 和 4%。实际上,这个假设并非事实,但上述百分比粗略估算了核苷酸水平上各种不同突变的相对频率 (Masatoshi and Sudhir, 2002)。本研究所得的核苷酸多样性各项分析数据表明,总体上同义位点的各项参数小于非同义位点,与上述假设大致相符,而且 CO1 基因发生突变的数目(突变数分别为 5 和 40),同义和非同义位点数目(33.67, 140.33)都远远小于 CO2 基因发生突变的数目(突变数分别为 100 和 80),同义和非同义位点数目(107.64, 444.36),推测 CO2 基因进化速率高于 CO1 基因进化速率。另外,本研究还发现密码子使用出现偏倚,这是因为:高表达基因的密码子使用频率与其细胞内的同功 tRNA 的丰度有关 (Ikemura,

1981, 1985)除了同功 tRNA 的相对丰度,还有一个其他因素影响密码子使用,这就是偏倚的突变压(biased mutation pressure)。当 $G + C$ (即 GC 含量)的相对频率发生变化时,人们认为这种变异主要是在核苷酸水平上 GC 至 AT 的正向和回复突变压的差异造成的(Sueoka, 1962)。当然,为了维持蛋白质的功能,由突变压决定的平衡频率也会同预期的密码子第 3 位的 GC 含量不同,因为某些密码子第 3 位的替代导致氨基酸置换而受到净化选择的淘汰作用;密码子第 2 位的替代,全部是非同义的,所以该替代主要由功能制约而不是突变压所制约;密码子第 1 位中有一小部分替代是同义的,因而突变压对其作用介于第 3 位和第 2 位之间。所以,密码子用法受突变压和净化选择两者双重控制。而且,位于不同同质区的基因可能有不同的密码子用法偏倚模式。由于偏倚模式对核苷酸替代速率的影响(Kusnezov, 1963; Wetterer *et al.*, 1998),这些基因将以不同速率发生演变。

通过排列的序列间转换与颠换的比率可以断定置换的饱和性,对新近分化的序列而言,在核基因组中转换与颠换的比率约等于 2.0,在动物的线粒体基因更高。本研究所得数据显示转换/颠换的比值接近 1.0 左右,可以认为本研究所用的基因序列的转换基本上达到了饱和。另外,CO2 基因转换/颠换的比值(1.040),比 CO1 基因转换/颠换的比值(0.929)更接近 1.0,说明 CO2 基因进化速率比 CO1 基因的进化速率快,因而比 CO2 基因先达到饱和状态,这也从另一个侧面说明了 CO1 基因相对与 CO2 基因是保守的。

Dowton 和 Austin(1999)通过对膜翅目 15 个科的广泛调查研究发现,在一些研究案例中,CO2 基因与 tRNA 基因之间,以及 tRNA 基因之间有重叠现象,而蚂蚁作为他们研究中的一个种类,却没有出现此现象。Dowton 和 Austin 对此的解释是:由于在他们所研究的 CO1 基因-tRNA_{Leu} 基因中,紧跟着 CO2 基因的 tRNA 基因被该 CO2 基因相对的那条链转录,从而降低了对多种腺苷进行酸化作用要求,同时也减少了对 tRNA 在第一次转录终止后精确脱落的需要。对于切叶蚁亚科来说,没有重叠的 tRNA 的精确脱落似乎是一种简约模式,假定如果没有发生多种腺苷酸化作用和 tRNA 的精确脱落,那么 CO1 的一些基因长度将会在很大程度上被延长。所以,基于上述解释,在切叶蚁亚科中观察到的这种翻译终止在进化过程中产生的分歧现象说明:单一个碱基

T 就可以满足该昆虫线粒体基因的翻译终止需要,同时,这种情况也同时发生在与该 T 碱基毗邻的 tRNA 基因中。

分析系统树构建结果显示,用 DNA 序列来建树,ML, NJ, UPGMA 方法所构建的分子系统树结果比较理想而且较一致,如果用氨基酸序列来进行构建分子系统树,对于 CO1 基因而言,其结果并不理想,不同系统树的分支之间较混乱。但是,如果仅分析 BCL 值较高的分支,其结果与用 DNA 序列分析所得到的结果大致相同;对于 CO2 基因,用氨基酸序列建树结果要比 CO1 基因显示的结果要好,但是仍存在许多矛盾的地方。同样,如果仅分析 BCL 值较高的分支,其结果与用 DNA 序列分析所得到的结果大致相同;可是,当把 CO1 基因与 CO2 基因的氨基酸序列合并起来再分析时,4 种方法所得到的结果比较理想,结果也比较一致,而且该结果也与用 DNA 序列进行分析所得到的结果一致。总体上看,全序列的分析结果较好,CO2 基因分析结果其次,CO1 基因分析结果最差,基于前面所提到的分析推测结果:CO1 基因比 CO2 基因保守。因此, DNA 序列片段保守程度这一因素很可能影响了系统发育的分析结果;另一方面,从与转录和翻译有关的真核蛋白编码基因的基本结构可以看出,从 DNA 到 mRNA 将会有大量的信息位点(如内含子, ICS, NCS 等)会被丢失。这一分析结果显示:在构建分子系统发育树并进行系统发育分析时,获得较多的位点信息是十分关键和必要的,如果位点信息太少,将很有可能得到不同的结论。用氨基酸构建分子系统发育树时,如果有效的信息位点太少,其分析所得到的结果差于用 DNA 序列分析所得到的结果。但是,随着信息位点的增加,用氨基酸序列进行分析所得到的结果将比用 DNA 序列分析所得到的结果越来越好,原因有以下几个方面:(1)氨基酸序列可以反应出同质结构,而核苷酸序列则不能;(2)如果把核苷酸序列作为物种间进化关系和系统发育研究的基础,那么,这种固有模式一旦被生物的趋同进化所打破,在关系较远的种类之间要比关系较近的种类之间也许将产生更多的相似点,从而导致了物种分类的依据可以是其基本组织结构而不是亲缘关系;(3)如果进化关系和系统发育分析是基于核苷酸序列则可能会产生很大的分歧,这也许是因为受到了核苷酸密码子位置之间相互关联的影响(Lockhart *et al.*, 1994; Adachi and Hasegawa, 1996)。

蛋白质的氨基酸序列重新被用于重建系统发

育,蛋白质 MP 法通常比 DNA 的 MP 法能获得更好的结果。另外,当 DNA 序列间关系较近时,DNA 似然法似乎较好,尤其是数据只使用第一和第二密码子时。但是,如果是距离较远的蛋白质编码基因,情况就很复杂,因为同义替代速率通常远高于非同义替代。位于第三密码子的 4 种核苷酸的相对频率也可随物种不同而变化。这说明核苷酸的替代的静止模型并不一定是合适的。相反,蛋白质序列的进化改变受上述问题的影响并不严重,当序列之间的差异较大时,它似乎比 DNA 序列研究更为简单 (Masatoshi and Sudhir, 2002)。

通过 ML、MP、NJ、UPGMA 4 种方法所构建的分子系统树,其结果都并不相同,这与这 4 种方法本身差异有关,也与每个方法所选择的取代模型相关。另外,自举检验计算,在合适的条件下(各种替换速率基本相等,树枝基本对称),也即自举置信水平大于 70%,那么所得到的系统发育进化树能够反映真实的系统发生史的可能性要大于 95%。同理,如果条件不是很合适,那么如果自引导数值大于 50%,则精确性的评估就会过高 (Hillis and Bull, 1993)。在某些条件下,如果自引导数值较高,可能会使系统发生史看起来很好,从而得出错误的结论。另外,本研究的取样数量和取样的不均一性对统计结果产生了一定影响,这个问题有待今后进一步研究解决。

本研究所用的 CO1 基因、CO2 基因以及 tRNA_{Leu} 基因应适合应用于属间进化关系及系统发育的分析,而并不适合于对种间,特别是对亲缘关系非常密切的物种之间的进化及系统发育关系进行分析研究和探讨。如果涉及的研究问题不仅仅是属间问题,而是包括了种间问题,这就需要选择和考虑并入更加合适的基因,如细胞色素 b 基因(Cytb),该基因片段进化速率要大于 CO1、CO2 基因的进化速率。这些问题的解决有待于今后深入研究。

此外,通过分析 tRNA_{Leu} 基因的二级结构获得了一些有意义的信息,这意味着研究分析由低级结构向高级结构发展,由基因结构到其行使的功能,由内部结构与机制再到表面性状与特征,这将是分子系统学今后发展的另一个趋势,正如当初传统的系统发育分析由宏观转入微观而产生了分子系统学,随着分子系统学研究的不断深入和发展,分子系统学应该,而且也会由微观向宏观分析研究方向发展,微观与宏观,此二者之间应该是相辅相成、相互促进的。

致谢 衷心感谢广西师范大学生命科学学院唐绍清教授给予研究上的指导,黄建华博士帮助收集大量的文献资料,彭云滔同学在实验过程中给予热心帮助。特别感谢中山大学张鹏博士、清华大学张静博士对本研究的实验和数据处理提出宝贵意见和建议。

参 考 文 献 (References)

- Adachi J, Hasegawa M, 1996. MOLPHY Version 2.3. Programs for Molecular Phylogenetics Based on Maximum Likelihood. Institute of Statistical Mathematics, Tokyo.
- Bourke AFG, Franks NR, 1995. Social Evolution in Ants. Princeton University Press, Princeton, NJ.
- Chiotis M, Jermiin LS, Crozier RH, 2000. A molecular framework for the phylogeny of the ant subfamily Dolichoderinae. *Mol. Phylogenet. Evol.*, 17: 108 - 116.
- Clary DO, Wolstenholme DR, 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. *J. Mol. Evol.*, 22: 252 - 271.
- Crozier RH, Crozier YC, 1993. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. *Genetics*, 133: 97 - 117.
- Crozier RH, Dobric N, Imai HT, Graur D, Cornuet JM, Taylor RW, 1995. Mitochondrial-DNA sequence evidence on the phylogeny of Australian jack-jumper ants of the *Myrmecia pilosula* complex. *Mol. Phylogenet. Evol.*, 4: 20 - 30.
- Crozier RH, Pamilo P, 1996. Evolution of Social Insect Colonies. Sex Allocation and Kin-Selection. Oxford Univ. Press, Oxford, UK.
- Crozier RH, Oldroyd BP, Tay WT, Kaufmann BE, Johnson RN, Carew ME, Jennings KM, 1997. Molecular advances in understanding social insect population structure. *Electrophoresis*, 18: 1 672 - 1 675.
- Dowton M, Austin AD, 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Mol. Biol. Evol.*, 16: 298 - 309.
- Fauron CMR, Wolstenholme DR, 1990. Intraspecific diversity of nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules of *Drosophila simulans*. *Nucleic Acids Res.*, 8: 5 391 - 5 410.
- Fitch WM, Margoliash E, 1967. Construction of phylogenetics trees. *Science*, 155: 279 - 284.
- Gojoberi T, Li WH, Graur D, 1982. Patterns of nucleotide substitutions in pseudogenes and functional genes. *J. Mol. Evol.*, 18: 360 - 369.
- Hillis DM, Bull JJ, 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.*, 42: 182 - 192.
- Hölldobler B, Wilson EO, 1990. The Ants. Harvard University Press, Cambridge, MA.
- Ikemura T, 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.*, 146: 1 - 21.
- Ikemura T, 1985. Codon usage and tRNA content in unicellular and

- multicellular organisms. *Mol. Biol. Evol.* , 2 : 13 - 34.
- Kocher TD , Wilson AC , 1991. Sequence evolution of mitochondrial DNA in humans and chimpanzees : Control region and a protein-coding region. In : Osawa S , Honjo T eds. *Evolution of Life : Fossils , Molecules and Culture*. Springer-Verlag , New York. 391 - 413.
- Kusnezov N , 1963. Zoogeografia de las hormigas en Sudamerica. *Acta Zool. Lilloana* , 19 : 3 - 186.
- Lockhart PJ , Steel MA , Hendy MD , Penny D , 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* , 11 : 605 - 612.
- Masatoshi N , Sudhir K , 2002. *Molecular Evolution and Phylogenetics*. Oxford University Press.
- Pashley DP , Ke LD , 1992. Sequence evolution in the mitochondrial ribosomal and ND-1 gene in Lepidoptera : implications for phylogenetic analyses. *Mol. Biol. Evol.* , 9 : 1 061 - 1 075.
- Qi YP , 1998. *Manipulation and Principle of Gene*. Wuhan : Wuhan University Press. 10 - 30. [齐义鹏 , 1998. *基因及其操作原理*. 武汉 : 武汉大学出版社. 10 - 30]
- Sueoka N , 1962. On the genetic basis of variation and heterogeneity of DNA base composition. *Proc. Natl. Acad. Sci. USA* , 48 : 582 - 592.
- Wetterer JK , Schultz TR , Meier R , 1998. Phylogeny of fungus-growing ants (tribe Attini) based on mtDNA sequence and morphology. *Mol. Phylogenet. Evol.* , 9 : 42 - 47.
- Wilson EO , 1992. Social insects as dominant organism. In : Bollen J ed. *Biology and Evolution of Social Insects*. Leuven : Belgium , Leuven University Press. 1 - 7.

(责任编辑 : 袁德成)