

云南松毛虫线粒体基因组全序列测定和分析

王瑶^{1,2}, 孔祥波², 张苏芳², 刘福², 张真², 严善春^{1*}

(1. 东北林业大学林学院, 森林生态系统可持续经营教育部重点实验室, 黑龙江 哈尔滨 150040; 2. 国家林业和草原局森林保护学重点实验室, 中国林业科学研究院森林生态环境与保护研究所, 北京 100091)

摘要: [目的] 测定和分析了云南松毛虫线粒体基因组特征, 从线粒体基因组水平探究鳞翅目蛾类昆虫高级阶元的系统发育关系。 [方法] 采用 Illumina HiSeq 测序方法测定了云南松毛虫线粒体全基因组序列, 参考鳞翅目昆虫已知线粒体基因组的全序列对其各基因进行定位和注释。采用 tRNA Scan-SE 2.0 在线预测云南松毛虫线粒体基因组 tRNA 基因的二级结构。基于线粒体全基因组的蛋白编码序列构建了鳞翅目 13 个科 32 种蛾类昆虫的系统发育树和松毛虫属近缘种间的系统发育树。 [结果] 结果显示云南松毛虫线粒体基因组全长 15 443 bp, 包括 13 个蛋白质编码基因, 22 个 tRNA 基因, 2 个 rRNA 基因和一段长度为 321 bp 的 A + T 富含区, 无基因重排, 存在较高的 A + T 含量 (80.0%)。13 个蛋白编码基因中除了 *ND2* 和 *COX1*, 其余均以 ATN 做为起始密码子。9 个蛋白编码基因共享相同的终止密码子 TAA (*ND2*, *ATP8*, *ATP6*, *COX3*, *ND5*, *ND4L*, *ND6*, *CYTB* 和 *ND1*), 其他 4 个基因的终止密码子都是残缺的, *COX1*, *COX2*, *ND4* 以 T 为终止密码子, *ND3* 以 TA 为终止密码子。22 个 tRNA 基因中, 除了 tRNA^{Ser(AGN)} 由于缺少 DHU 臂无法构成三叶草结构, 其余 T 均为典型的三叶草结构。整个线粒体结构与鳞翅目中目前已得到的其他昆虫线粒体基因组结构一致。 [结论] 系统发育分析结果显示, 云南松毛虫与思茅松毛虫是完全不同的近缘种, 与其他松毛虫亲缘关系也较远, 云南松毛虫与 6 种近缘种的系统发育关系为: (((((油松毛虫 + 文山松毛虫) + 赤松毛虫) + 落叶松毛虫) + (思茅松毛虫 + 云南松毛虫)) + 家蚕)。鳞翅目蛾类各科之间的系统发育关系为: (((((((毒蛾科 + 灯蛾科) + 夜蛾科) + 舟蛾科) + (尺蛾科 + (蚕蛾科 + (天蛾科 + 大蚕蛾科)))) + 枯叶蛾科) + (螟蛾科 + 草螟科)) + 卷蛾科) + 蝙蝠蛾科)。

关键词: 鳞翅目; 枯叶蛾科; 云南松毛虫; 线粒体基因组; 系统发育

中图分类号: S763.3

文献标识码: A

文章编号: 1001-1498(2019)05-0011-10

云南松毛虫 (*Dendrolimus houi* Lajonquiere) 又名大柏毛虫, 属鳞翅目 (Lepidoptera) 枯叶蛾科 (Lasio-campidae) 松毛虫属 (*Dendrolimus*), 主要危害柏木、柳杉、思茅松和马尾松等针叶树种^[1]。该虫严重威胁森林生态安全及林业可持续发展^[2], 因此我国对于其生物学特性、防治和管理等方面都有了较为深入的研究。孔祥波^[3]等人鉴定了云南松毛虫的性信息素包括 E5, Z7-12:OH 和 E5, Z7-12:OAc 两种成分, 成功应用于林间种群动态监测。关于松毛虫属内近缘种之间亲缘关系的研究也获得了较大成果。2004—2015 年期间多种分子标记被应用于松毛虫属内亲缘关系的研究^[4-10], 综合不同的分子标记获得了一个相似的结果, 云南松毛虫与思茅松毛虫亲

缘关系较近, 马尾松毛虫、文山松毛虫、德昌松毛虫、油松松毛虫、赤松毛虫和落叶松毛虫亲缘关系相对较近。

线粒体基因组 (mtDNA) 是一种双链闭环状的 DNA 分子, 在鳞翅目昆虫体内编码结构稳定, 可操作性强, 进化速率快^[11-12], 近年来被广泛应用于解释物种间的系统发育关系^[13-14]。松毛虫属近缘种昆虫 (马尾松毛虫 NC_027156.1, 赤松毛虫 KJ_913815, 油松毛虫 KJ_913817, 思茅松毛虫 MF100138.1) 已经完成了线粒体基因组的测序, 但并未见云南松毛虫的报道。本研究通过高通量测序获得云南松毛虫的线粒体全基因组序列, 并结合已报道的鳞翅目蛾类的线粒体基因组进行比较分析,

收稿日期: 2018-11-28 修回日期: 2019-02-12

基金项目: 国家自然科学基金项目 (31470654)

* 通讯作者: E-mail: yanshanchun@126.com

探究松毛虫所在的高级阶元的分类地位,以期丰富云南松毛虫线粒体全基因组的基础数据,为松毛虫属系统学和分类学的研究提供更准确的分子生物学数据。

1 材料与方法

1.1 供试虫源

云南松毛虫于2016年9月采自云南省景东县(23°56'N,101°15'E),采集到的活体昆虫固定在无乙醇中,置-20℃冰箱保存备用。样本由中国林业科学研究院孔祥波老师鉴定。

1.2 Illumina Hiseq 高通量测序及基因注释

上海美吉生物公司对云南松毛虫样品本进行

DNA提取,使用Covaris M220超声破碎,DNA片段化为400~500 bp。制备DNA文库,扩增后将符合要求的文库安排Illumina Hiseq仪器进行paired-end(PE)测序。将获得的原始序列进行质控,利用SOAPdenovo v2.04拼接软件对质控后的数据进行de novo组装,将组装得到的contig与线粒体数据库进行比对,筛选得到来自线粒体的contig;然后利用MITObim v1.6迭代比对,将测序的所有clean reads mapping于contig上延伸,通过gap close获得线粒体的全基因组序列。利用DOGMA软件对基因组中包含的gene、rRNA和tRNA进行预测。确定无误后上传至NCBI数据库。

表1 用于高级阶元系统发育分析的物种信息

Table 1 Information of species used for phylogenetic analysis

科名 Family	物种 Species	线粒体基因组大小 mt DNA Size/bp	(A + T)/%	GenBank 登录号 GenBank accession no.
枯叶蛾科 Lasiocampidae	云南松毛虫 <i>Dendrolimus houi</i>	15 443	80.0	-
	马尾松毛虫 <i>Dendrolimus punctatus</i>	15 417	79.5	NC_027156.1
	思茅松毛虫 <i>Dendrolimus kikuchii</i>	15 422	79.2	MF100138.1
	油松毛虫 <i>Dendrolimus tabulaeformis</i>	15 411	79.5	KJ_913817
	赤松毛虫 <i>Dendrolimus spectabilis</i>	15 412	79.3	KJ_913815
	落叶松毛虫 <i>Dendrolimus superans</i>	15 417	80.1	NC_039841.1
	竹纹枯叶蛾 <i>Euthrix laeta</i>	15 368	80.2	KU870700.1
	青黄枯叶蛾 <i>Trabala vishnou guttata</i>	15 281	80.9	KU884483.1
天蛾科 Sphingidae	烟草天蛾 <i>Manduca sexta</i>	15 516	81.8	NC_010266.1
	<i>Sphinx morio</i>	15 299	81.1	NC_020780.1
大蚕蛾科 Saturniidae	天蚕 <i>Samia cynthia ricini</i>	15 384	79.7	NC_017869.1
	章蚕 <i>Eriogyna pyretorum</i>	15 327	80.8	NC_012727.1
夜蛾科 Noctuidae	小地老虎 <i>Agrotis ipsilon</i>	15 377	81.3	KF163965.1
	草地夜蛾 <i>Spodoptera frugiperda</i>	15 365	81.1	KM362176.1
	斜纹夜蛾 <i>Spodoptera litura</i>	15 383	81.1	KF701043
毒蛾科 Lymantriidae	舞毒蛾 <i>Lymantria dispar</i>	15 569	79.9	FJ617240.1
	黄尾毒蛾 <i>Euproctis similis</i>	15 437	80.2	NC_035886.1
	茶毒蛾 <i>Euproctis pseudoconsersa</i>	15 461	79.7	KJ716847.1
尺蛾科 Geometridae	桑尺蠖 <i>Phthonandria atrilineata</i>	15 499	81.0	NC_010522.1
蝙蝠蛾科 Hepialidae	<i>Ahamus yunnanensis</i>	15 816	82.3	NC_018095.1
	人支蝠蛾 <i>Thitarodes renzhiensis</i>	16 173	81.3	NC_018094.1
蚕蛾科 Bombycidae	桑蚕 <i>Rhodotia menciaria</i>	15 306	78.8	NC_021962.1
	家蚕 <i>Bombyx mori</i>	15 643	81.3	NC_002355.1
	野桑蚕 <i>Bombyx mandarina</i>	15 928	81.7	NC_003395.1
灯蛾科 Arctiidae	美国白蛾 <i>Hyphantria cunea</i>	15 481	80.4	NC_014058.1
螟蛾科 Pyralidae	亚洲玉米螟 <i>Ostrinia furnacalis</i>	14 563	80.4	NC_003368.1
舟蛾科 Notodontidae	苹掌舟蛾 <i>Phalera flavescens</i>	15 659	80.9	NC_016067.1
	杨扇舟蛾 <i>Clostera anachoreta</i>	15 456	80.7	NC_034740.1
草螟科 Crambidae	小蔗秆草螟 <i>Diatraea saccharalis</i>	15 490	80.1	NC_013274.1
	二化螟 <i>Chilo suppressalis</i>	15 395	80.6	NC_015612.1
卷蛾科 Tortricidae	茶小卷叶蛾 <i>Adoxophyes honmai</i>	15 680	80.4	DQ073916.1
	梨小食心虫 <i>Grapholita molesta</i>	15 717	80.9	HQ392511.1
双翅目 Diptera	地中海实蝇 <i>Ceratitidis capitata</i>	15 980	77.5	AJ242872.1
实蝇科 Tephritidae	桃果实蝇 <i>Bactrocera zonata</i>	15 935	73.4	NC_027725.1

表2 云南松毛虫线粒体基因组注释

Table 2 Organization of the mitochondrial genome of *Dendrolimus houi*

基因 Gene	方向 Direction	位置 Location	长度 Size	反密 码子 Anticodon	起始密 码子 Start codon	终止密 码子 Stop codon	基因 间隔 Inter- genic
trnM	F	1	67	67	CAT		
trnI	F	69	132	64	GAT		7
trnQ	R	130	198	69	TTG		-1
ND2	F	257	1 264	1 008		TTT TAA	58
trnW	F	1 263	1 332	70	TCA		-2
trnC	R	1 325	1 391	67	GCA		-8
trnY	R	1 393	1 463	71	GTA		0
COX1	F	1 492	3 022	1 531		CGA T	34
trnL	F	3 023	3 089	67	TAA		0
COX2	F	3 090	3 771	682		ATA T	0
trnK	F	3 775	3 845	71	CTT		0
trnD	F	3 847	3 913	67	GTC		3
ATP8	F	3 914	4 075	162		ATT TAA	0
ATP6	F	4 069	4 746	678		ATG TAA	-7
COX3	F	4 762	5 550	789		ATG TAA	14
trnG	F	5 553	5 618	66	TCC		2
ND3	F	5 619	5 971	353		ATT TA	0
trnA	F	5 971	6 038	68	TGC		-1
trnR	F	6 048	6 112	65	TGG		15
trnN	F	6 118	6 184	67	GTT		4
trnS1	F	6 211	6 276	66	GCT		11
trnE	F	6 277	6 345	69	TTC		-1
trnF	R	6 354	6 420	67	GAA		4
ND5	R	6 425	8 167	1 743		ATT TAA	2
trnH	R	8 168	8 231	64	GTG		0
ND4	R	8 232	9 570	1 339		ATG T	0
ND4L	R	9 602	9 895	294		ATG TAA	24
trnT	F	9 900	9 964	65	TGT		7
trnP	R	9 965	10 029	65	TGG		0
ND6	F	10 038	10 568	531		ATA TAA	8
CytB	F	10 574	11 722	1 149		ATG TAA	4
trnS2	F	11 726	11 791	66	TGA		3
ND1	R	11 791	12 744	954		ATG TAA	-1
trnL	R	12 746	12 813	68	TAG		1
rrnL	R	12 840	14 278	1 439			0
trnV	R	14 284	14 349	66	TAC		0
rrnS	R	14 350	15 122	773			0
AT region	F	15 123	15 443	321			0

基因间隔一列中,正数表示基因间隔碱基数,负数表示基因重叠碱基数。

In the column of intergenic length, the positive number indicates interval base pairs between genes, while the negative indicates the overlapping base pairs between genes.

2 结果与分析

2.1 云南松毛虫线粒体基因组全序列分析

云南松毛虫线粒体基因组全序列总长度为

15 443 bp,呈双链闭环环状结构,基因组成包含 13 个蛋白质编码基因(PCGs)、22 个 tRNA 基因、2 个 rRNA 基因(12S rRNA 和 16SrRNA),1 个 A + T 富含区(A + T-rich region)(如图 1)。详细注释见表 2。37 个基因相邻之间存在基因间隔或基因重叠现象。云南松毛虫线粒体全基因组序列中 A、T、C、G 的碱基含量分别为 41.2%、38.8%、12.4% 和 7.6% (表 3),表现出明显的 AT 偏好性。基因组全序列碱基 AT-skew 值为 0.03%,GC-skew 值为 -0.24%,tRNA 基因的 AT 偏斜和 GC 偏斜均为正值,rRNA 基因的 AT 偏斜为负,GC 偏斜为正。

表3 云南松毛虫线粒体基因组核苷酸组成

Table 3 Nucleotide composition of the complete mitochondrial genome of *Dendrolimus houi*

基因序列 Gene sequence	Size /bp	A/%	T/%	C/%	G/%	A + T /%	C + G/%
全基因组 Whole genome	15 443	41.2	38.8	12.4	7.6	80.0	20.0
蛋白编码 基因 PCGs	11 213	33.9	44.4	10.6	11.1	78.3	21.7
rRNA 基因	1 439	40.4	44.1	4.4	11.1	84.5	15.5
tRNA 基因	773	44.1	41.5	4.6	9.8	85.6	14.4
A + T 富集区	321	40.8	51.4	5.0	2.8	92.2	7.8

2.2 云南松毛虫线粒体基因组蛋白编码基因

云南松毛虫线粒体基因组共编码 13 个蛋白质基因(*COX1*, *COX2*, *COX3*, *ATP6*, *ATP8*, *ND1*, *ND2*, *ND3*, *ND4*, *ND4L*, *ND5*, *ND6*, *Cyt B*)。除了 *COX1* 以 CGA 为起始密码子,其余均以 ATN 作为起始密码子。9 个蛋白编码基因共享相同的终止密码子 TAA (*ND2*, *ATP8*, *ATP6*, *COX3*, *ND5*, *ND4L*, *ND6*, *CYTB* 和 *ND1*),其他 4 个基因的终止密码子都是残缺的,*COX1*、*COX2*、*ND4* 以 T 为终止密码子,*ND3* 以 TA 为终止密码子。这种不完整的终止密码子在鳞翅目线粒体基因组中是很常见的,有研究推测它们会在转录后的多腺苷酸化过程中变成完整体^[16]。云南松毛虫线粒体基因组的全部蛋白质编码基因的氨基酸使用情况(图 2:A)和相对密码子使用率 RSCU(图 2:B)的分析结果显示,云南松毛虫 13 个蛋白质编码基因中氨基酸使用最频繁的依次包括 Phe、Ile、Leu 和 Tyr(Count > 300),而 Ala、Met 和 Trp 的使用相对较少(Count < 50);相对同义密码子中 UUA 使用频率最高,RSCU 值为 3.12,密码子 CGC 和 CCG 的相对密码子使用频率最低,其 RSCU 值分别为 0 和 0.07。

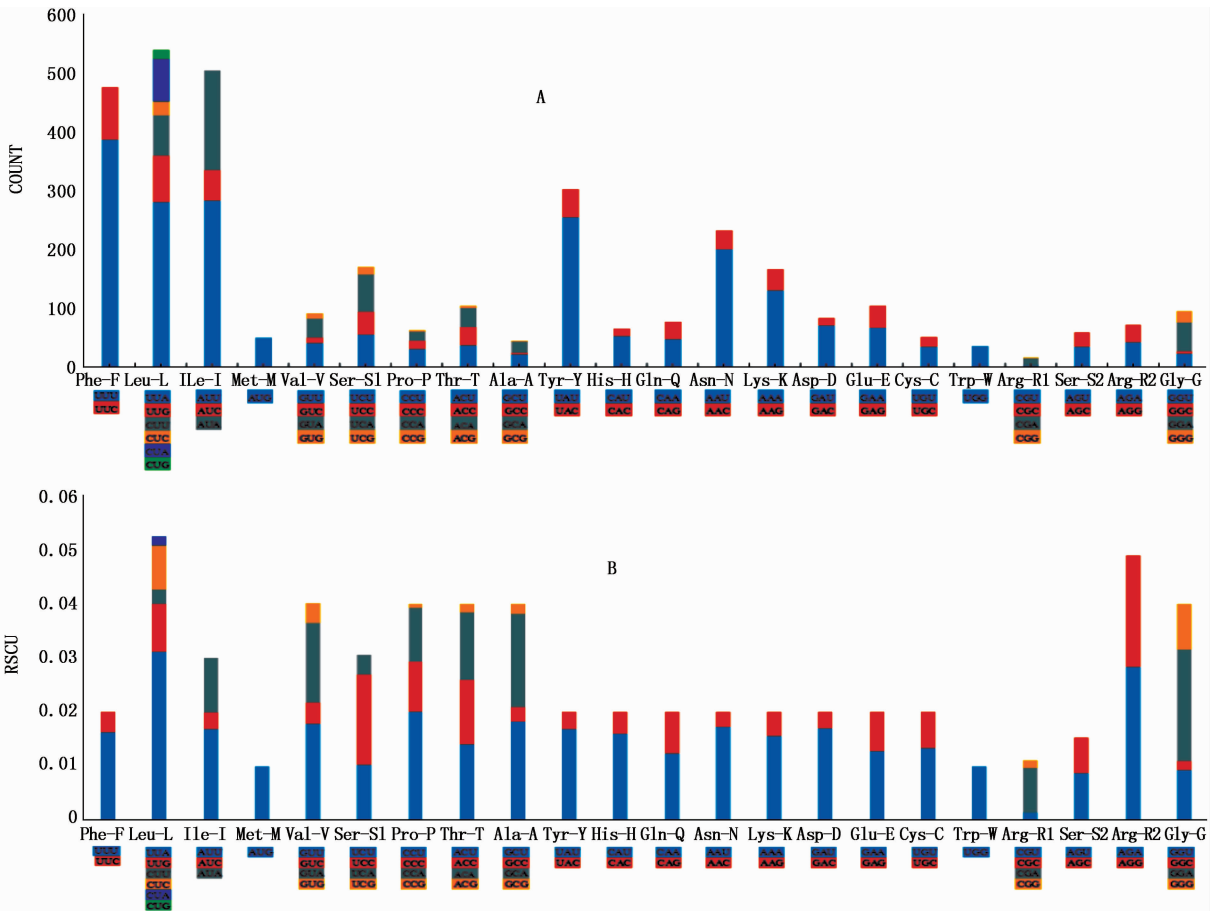


图 2 云南松毛虫基因组蛋白质编码基因氨基酸的使用频率 (A) 和同义密码子相对使用度 (B)

Table 2 Usage of amino acids (A) and the relative synonymous codon usage (B) of protein-coding genes of mitochondrial genome in *Dendrolimus houi*

2.3 tRNA 基因及 rRNA 基因的结构分析

云南松毛虫线粒体基因组上共检测出 22 个 tRNA 基因, A + T 的含量为 81.3%, 总长为 1 475 bp, 长度范围为 64 ~ 71 bp, 在线粒体基因组上的排列顺序与其它松毛虫属昆虫相同。其中 14 个编码在 J 链上 (major coding strand), 8 个编码在 N 链上 (minor coding strand)。所有的 tRNA 都有经典的三叶草二级结构 (图 3), 除了 trnS1 (AGN) 基因, 它的二氢尿苷臂 (DHU) 被未配对的核苷酸取代, 无法形成典型的三叶草结构。在云南松毛虫线粒体基因组的 22 个 trnA 基因的二级结构中出现了 22 对碱基错配现象。其中 G-U 错配 20 个, U-C 和 U-U 错配各 1 个。U-C 错配出现在 tRNA^{Thr(TGT)} 的 TΨC 臂上, U-U 错配位于 tRNA^{Leu(TAA)} 的氨基酸接受臂上。20 个 G-U 错配在氨基酸接受臂、TΨC 臂、二氢尿嘧啶臂和反密码子臂上均有出现。这种错配普遍存在于鳞翅目昆虫的 tRNA 二级结构中^[18]。这些错配是符合摆

动配对原则的, 对维持 tRNA 二级结构的稳定起着重要的作用。

rRNA 基因是一类相对分子量较大的 RNA 基因, 可与蛋白质结合形成核糖体。云南松毛虫线粒体的核糖体小亚基 rrmS 基因和核糖体大亚基 rrmL 基因的长度分别是 773 bp 和 1 439 bp, 分别位于 A + T 富含区与 trnA-Val (TAC) 之间和 trnA-Val (TAC) 和 trnA-Leu (TAG) 之间。rrmS 基因和 rrmL 基因的 A + T 含量为 84.9% (表 3), 具有明显 AT 碱基偏向性。

2.4 A + T 富含区的结构分析

A + T 富含区为线粒体基因组中的非编码序列, 即为不能转录为相应的 mRNA, 从而不能指导合成相应蛋白质的区域。云南松毛虫线粒体基因组的 A + T 富含区位于 rrmS 基因与 trnA-Met 基因之间, 其长度为 321 bp, A + T 含量为 92.2%。rrmS 基因下游有一段由“ATAGA”引导的长度为 14 bp 的多聚 T

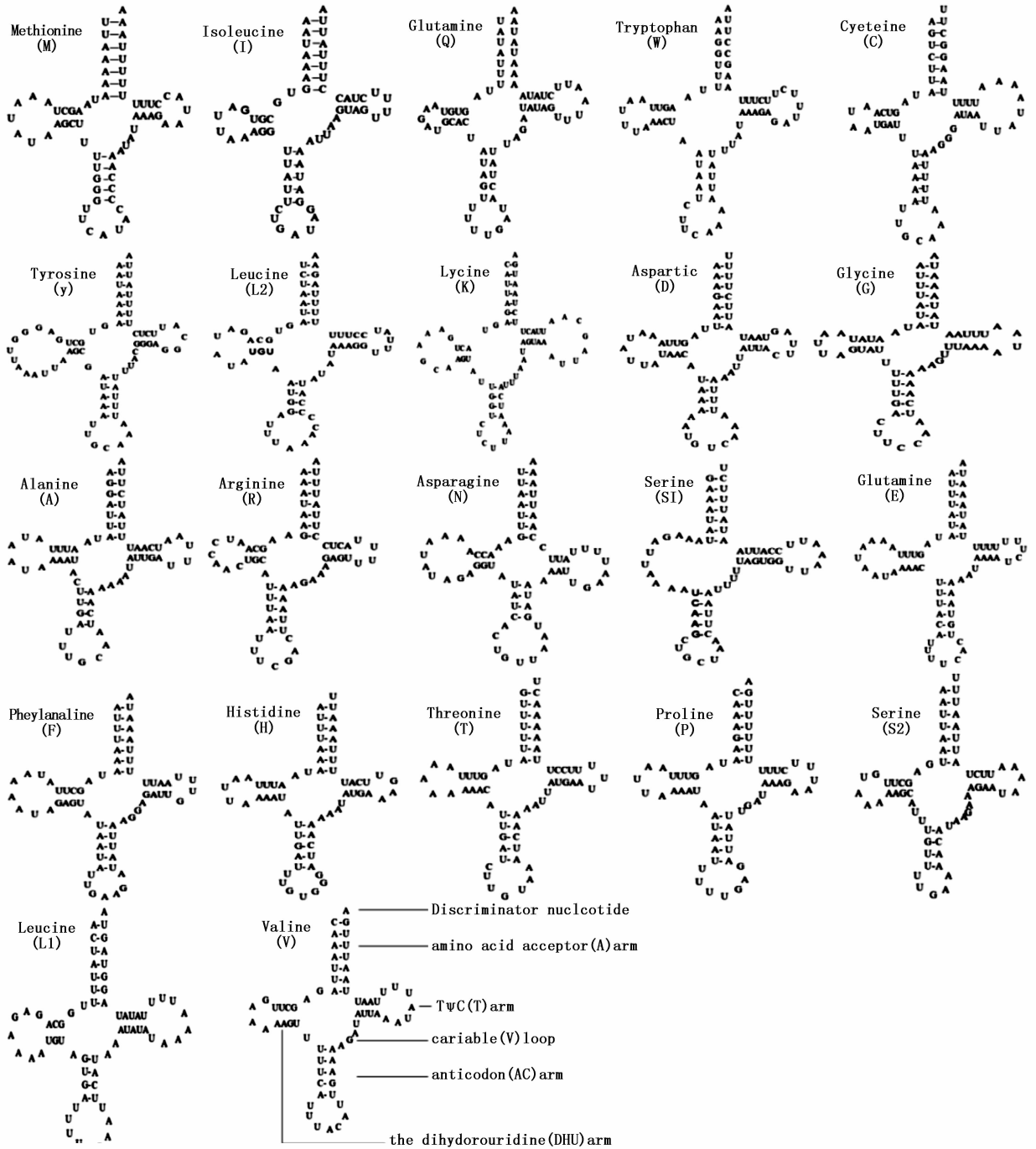


图3 云南松毛虫 tRNA 的二级结构

Fig. 3 Secondary structures for Trna genes from *Dendrolimus houi* mtDNA

结构(图4),这段多聚 T 结构在所有已测定的鳞翅目昆虫 mtDNA 中都很保守,长度从 12 bp(宽尾凤蝶 *Aeghana maraho*)到 22 bp(野桑蚕 *Bombyx mandarina*)不等。在云南松毛虫线粒体 A + T 富含区中还存在一个(AT)₇ 结构,这在其他鳞翅目昆虫中均有检测到类似结构^[13-14]。

2.5 云南松毛虫高级阶元系统发育关系

以双翅目 Diptera 实蝇科 Tephritidae 地中海实蝇 *Ceratitis capitata* 和桃果实蝇 *Bactrocera zonata* 为外群,选用已测定的 13 个科的鳞翅目蛾类昆虫和本研究中的云南松毛虫线粒体全基因组,基于 13 个蛋白编码基因的核苷酸序列构建 ML 树。结果显示

```
rms--15122bp
TATTTAATATATACATTTTTTAATATAGATTTTTTTTTTTTTTAT
polyT (14)
ATTAAATATTTATAGACATTATTAATTTTTTAATAATTTCTCT
TTTTTTTTTTCATAATATTCATATTAATACAAATGGCTATATA
AAATTTGACAATTTGATAAATTACAAAATAATAATATAATTAAT
ATTAATTTTTTCAGTTATTTATATATTAATATATATATATACAA
Microsatellite (AT) n element
TAAATAATTTAATATATTAATTAATTTTTAATAATAATAATTAAT
TTAAATTATTTGGAACCATTTGCAATTTTTTACTAATAATTA
ATTA
tRNA-Met
```

图 4 云南松毛虫 A + T 富含区结构

Fig. 4 Structure of (A + T)-rich region of *Dendrolimus houii* mitochondrial genome

(图 5), 枯叶蛾科中云南松毛虫与其他物种的系统发育关系为:((((((马尾松毛虫 + 油松毛虫) + 赤松毛虫) + 落叶松毛虫) + (思茅松毛虫 + 云南松毛虫)) + 竹纹枯叶蛾) + 青黄枯叶蛾)。13 种鳞翅目昆虫的系统发育分析结果显示, 枯叶蛾科 Lasiocampidae 与蚕蛾科 Bombycidae、大蚕蛾科 Satumiidae、天蛾科 Sphingidae 和尺蛾科 Geometriidae 聚为一簇, 并且有很高的支持率; 来自夜蛾总科的舟蛾科 Notodontidae, 夜蛾科 Noctuidae, 灯蛾科 Hypercompe 和

毒蛾科 Lymantriidae 聚为一簇, 系统发育关系为: ((毒蛾科 + 灯蛾科) + 夜蛾科 + 舟蛾科); 螟蛾总科的螟蛾科 Pyralidae 与草螟科 Cramnidae 聚为一簇; 卷蛾科 Tortricidae 和蝙蝠蛾科 Hepialidae 单独聚类。

2.6 松毛虫属内系统发育关系

松毛虫属内系统发育结果显示(图 6), 松毛虫属几个近缘种以较高的支持率分为两支。系统发育关系为: (((((油松毛虫 + 文山松毛虫) + 赤松毛虫) + 落叶松毛虫) + (思茅松毛虫 + 云南松毛虫)) + 家蚕)。基于 K2P 模型计算松毛虫属内种间的遗传距离(表 4), 结果显示: 思茅松毛虫与其他松毛虫的遗传距离最大。马尾松毛虫、文山松毛虫、赤松毛虫、油松毛虫与云南松毛虫间的遗传距离略小于思茅松毛虫与云南松毛虫的遗传距离; 油松毛虫与马尾松毛虫间的遗传距离(0.011 8) 小于文山松毛虫与马尾松毛虫间(0.013 5) 的遗传距离, 三者与赤松毛虫间的遗传距离均较大, 表明赤松毛虫与三者亲缘关系较远。

表 4 松毛虫属间的 K2P 遗传距离

Table 4 K2P genetic distance between species of the *Dendrolimus*

种群代码 population code	<i>D. pun</i>	<i>D. pun ws</i>	<i>D. tab</i>	<i>D. spet</i>	<i>D. sup</i>	<i>D. kiku</i>	<i>D. houii</i>	<i>Bombyx mori</i>
<i>D. pun</i>								
<i>D. pun ws</i>	0.013 5							
<i>D. tab</i>	0.011 8	0.007 1						
<i>D. spet</i>	0.051 2	0.052 2	0.051 8					
<i>D. sup</i>	0.070 6	0.072 7	0.071 9	0.071 0				
<i>D. kiku</i>	0.098 2	0.100 3	0.098 9	0.100 4	0.097 5			
<i>D. houii</i>	0.092 5	0.095 3	0.093 4	0.091 2	0.086 8	0.094 2		
<i>Bombyx mori</i>	0.193 6	0.195 6	0.194 5	0.193 5	0.191 7	0.198 5	0.189 4	

3 讨论

云南松毛虫线粒体基因组全长 15 443 bp, A + T 含量为 80.0%, 基因组的排列同已报道的鳞翅目其他昆虫的线粒体基因排列一致^[17-18]。其中 trnM-trnI-trnQ 的排列方式与假想昆虫祖先的排列方式(trnI-trnQ-trnM)不同^[19], 但是这种排列在鳞翅目昆虫中非常典型。鳞翅目线粒体基因组进化上的保守性, 表明其适合解决科、目等高级阶元的系统发育问题^[20]。

松毛虫属内系统发育树图显示, 云南松毛虫和思茅松毛虫有共同祖先, 而思茅松毛虫较早的与云南松毛虫分化, 二者独立进化。有研究表明, 二者的触角转录组表现出高度相似性, 证明它们有很近的

亲缘关系。从信息素角度出发, 思茅松毛虫性信息素成分包括 Z5, E7-12:OAc、Z5, E7-12:OH 和 Z5-12:OAc 3 种成分^[21], 而云南松毛虫性信息素成分包括 E5, Z7-12:OH 和 E5, Z7-12:OAc^[22], 不同的性信息素成分使它们产生了生殖隔离。马-文-油-赤拥有共同祖先, 且分化的时间较晚。云南松毛虫与思茅松毛虫的共同祖先种群与落-赤-马-油-文 5 种松毛虫的共同祖先互为姊妹群, 拥有更高级的共同祖先。从松毛虫的分化情况来看, 云南松毛虫应该是松毛虫属的基础物种, 与其他松毛虫物种亲缘关系较远, 这与 Alexander Kononov 等人^[23] 基于线粒体 *COI*、*COII* 和核基因 ITS2 的研究结果一致, 遗传距离结果也显示云南松毛虫与其他松毛虫之间存在较大的遗传距离。遗传距离结果中油松毛虫 × 文山

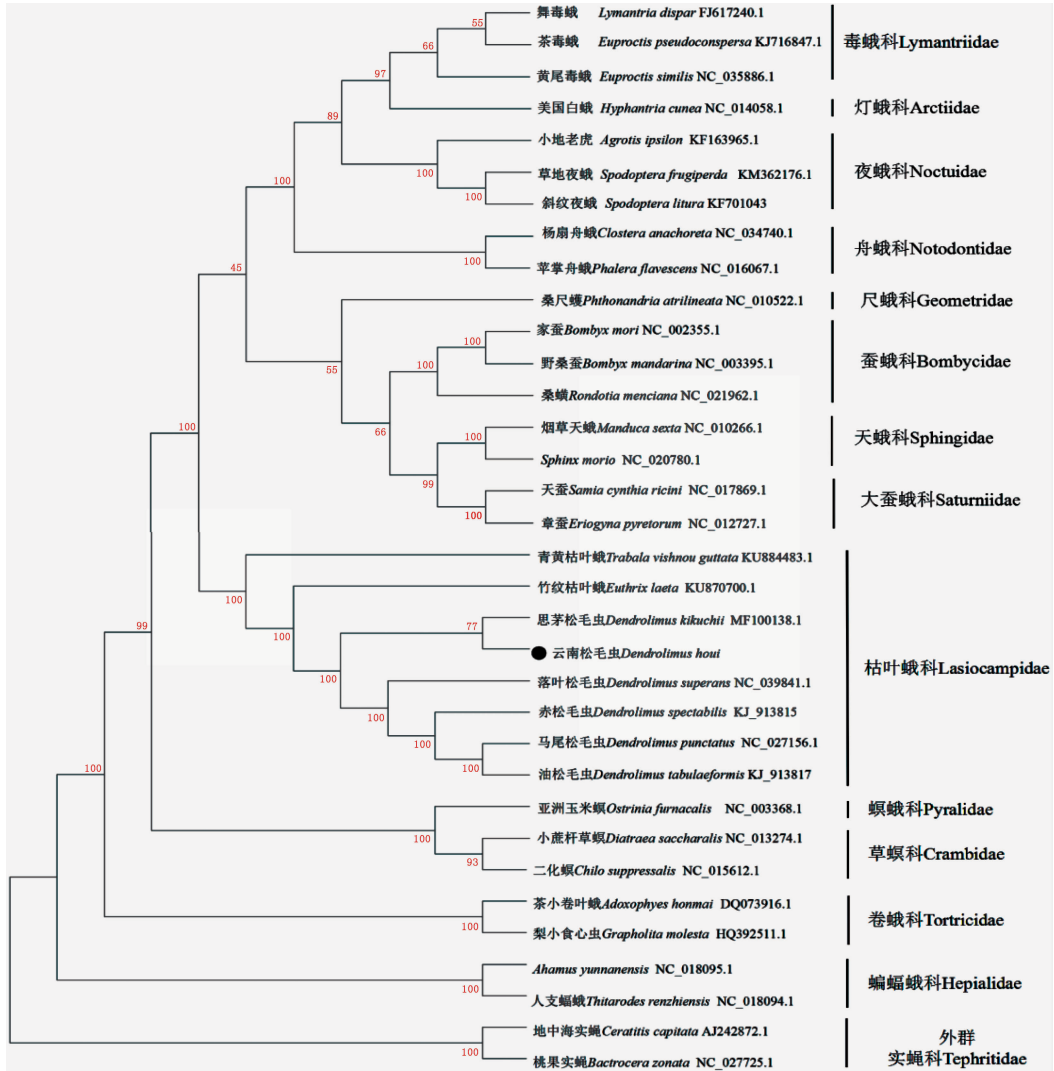


图 5 基于 32 种鳞翅目蛾类昆虫线粒体蛋白编码基因的最大似然树

Fig. 5 Maximum likelihood tree based on 32 mitochondrial moth insect mitochondrial PCGS

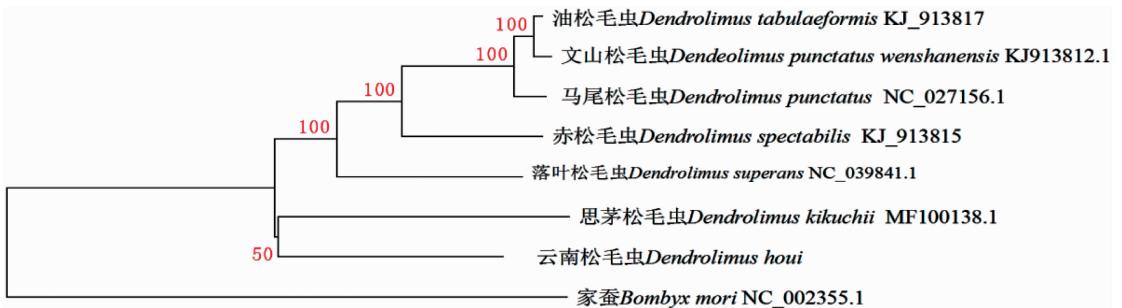


图 6 基于松毛虫属 13 个蛋白编码基因的 NJ 系统发育树

Fig. 6 NJ phylogenetic tree based on 13 protein-coding genes of *Dendrolimus*

松毛虫的遗传距离(0.007 1) < 油松毛虫 × 马尾松毛虫间的遗传距离(0.011 8) < 文山松毛虫 × 马尾松毛虫间(0.013 5)的遗传距离,而文山松毛虫是马

尾松毛虫公认的地理亚种^[10],这也支持油松毛虫和文山松毛虫作为马尾松毛虫地理亚种的分类地位。

鳞翅目蛾类高级分类阶元系统发育结果显示,枯叶蛾科与蚕蛾总科、尺蛾总科和夜蛾总科聚为一

支,其系统发育关系为:(枯叶蛾科+(夜蛾总科+(尺蛾总科+蚕蛾总科))),这与 Qin 等人^[10]报道的枯叶蛾科、蚕蛾总科和尺蛾总科的系统发育关系:(((天蛾科+蚕蛾科)+大蚕蛾科+尺蛾科)+枯叶蛾科)相一致。夜蛾总科中的舟蛾科、夜蛾科、毒蛾科和灯蛾科的进化关系为:(舟蛾科+(夜蛾科+(毒蛾科+灯蛾科))),舟蛾科、夜蛾科和毒蛾科的进化关系和吴榴宇^[24]得到的科级阶元系统发育结果:(裳蛾科+毒蛾科)+夜蛾科)+舟蛾科相一致。由此可见鳞翅目蛾类几大总科之间的系统发育关系较稳定,该结果在多数研究中均被验证^[25-26]。本研究通过测定云南松毛虫线粒体全基因组序列,分析其碱基组成、结构特点及其在高级阶元下与其他物种间的系统发育关系,丰富了枯叶蛾科线粒体基因组数据库,为枯叶蛾科物种鉴定、分类和进化分析奠定了分子基础。

4 结论

测定和分析了云南松毛虫线粒体基因组特征,全长 15 443 bp,包括 13 个蛋白质编码基因,22 个 tRNA 基因,2 个 rRNA 基因和一段长度为 321bp 的 A+T 富含区,无基因重排,存在较高的 A+T 含量(80.0%)。整个线粒体结构与鳞翅目中目前已得到的其他昆虫线粒体基因组结构一致。获得了松毛虫属内近源种间系统发育关系:((((油松毛虫+文山松毛虫)+赤松毛虫)+落叶松毛虫)+(思茅松毛虫+云南松毛虫))+家蚕;鳞翅目蛾类各科之间的系统发育关系:(((((((毒蛾科+灯蛾科)+夜蛾科)+舟蛾科)+(尺蛾科+(蚕蛾科+(天蛾科+大蚕蛾科))))+枯叶蛾科)+(螟蛾科+草螟科))+卷蛾科)+蝙蝠蛾科)。

参考文献:

[1] 侯陶谦. 中国松毛虫[M]. 北京:科学出版社,1987.6-31.
 [2] 陈昌洁. 松毛虫综合管理[M]. 北京:中国林业出版社,1990.5-25.
 [3] Kong X B, Zhang Z, Zhao C H, et al. Female Sex Pheromone of the Yunnan Pine Caterpillar Moth *Dendrolimus houi*: First (E,Z)-Isomers in Pheromone Components of *Dendrolimus* spp[J]. Journal of Chemical Ecology, 2007, 33(7): 1316-1327.
 [4] 张爱兵,孔祥波,李典谟,等. 中国松毛虫属八个种和亚种亲缘关系的 DNA 指纹证据[J]. 昆虫学报,2004(2): 236-242.
 [5] 南宫自艳,高宝嘉,杨 君. 松毛虫属(*Dendrolimus*) 部分种类亲缘关系与遗传分化的等位酶分析[J]. 生态学报,2009, 29(4): 1661-1667.

[6] 南宫自艳,杨 君,徐秀平,等. 四种松毛虫不同地理种群遗传多样性的等位酶分析[J]. 昆虫学报,2008(4): 417-423.
 [7] 南宫自艳,高宝嘉,徐志娥,等. 松毛虫属 5 种松毛虫不同地理种群 RAPD 遗传多样性分析[J]. 农业生物技术学报,2009, 17(1): 178-179.
 [8] 高宝嘉,高立杰,侯建华,等. 三种松毛虫不同地理种群遗传多样性[J]. 生态学报,2008(2): 842-848.
 [9] 贾玉迪,孔祥波,张 真,等. 基于线粒体 COI 基因部分序列分析 4 种松毛虫的亲缘关系[J]. 东北林业大学学报,2011, 39(11): 67-70+78.
 [10] Qin J, Zhang Y Z, Kong X B, et al. Mitochondrial phylogenomics and genetic relationship of closely related pine moth (*Lasiocampidae*: *Dendrolimus*) species in China, using whole mitochondrial genomes[J]. BMC Genomics, 2015,16(1): 428-439.
 [11] 张亚平,施立明. 动物线粒体 DNA 多态性的研究概况[J]. 动物学研究,1992(3): 289-298.
 [12] Brown W M, George M, Wilson A C. Rapid evolution of animal mitochondrial DNA[J]. Proceedings of the National Academy of Sciences, 1979, 76(4): 1967-1971.
 [13] Kong W, Yang J. The complete mitochondrial genome of *Rondotia menciiana* (Lepidoptera: Bombycidae)[J]. Journal of Insect Science, 2015, 15(1):48-48.
 [14] 张 敏,赵 盼,尹 洁,等. 小红珠绢蝶线粒体基因组特征及基于线粒体基因组的蝶类高级阶元系统发育关系分析[J]. 昆虫学报,2017, 60(11): 1324-1338.
 [15] Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets[J]. Molecular Biology & Evolution, 2016, 33(7):1870.
 [16] Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria [J]. Nature, 1981, 290(5806): 470-474.
 [17] Liu J H, Fu J Q, Yang L Y, et al. Complete mitochondrial genome of Simao pine caterpillar moth, *Dendrolimus kikuchii* (Lepidoptera: Lasiocampidae) and the related phylogenetic analysis[J]. Mitochondrial DNA Part B:Resources,2017,2(2),611-612.
 [18] 王菊平,曹天文,张 越,等. 扬眉线蛱蝶线粒体基因组全序列测定和分析[J]. 昆虫学报,2017, 60(8): 950-961.
 [19] Boore J L, Lavrov D V, Brown W M. Gene translocation links insects and crustaceans [J]. Nature, 1998, 392 (6677): 667-668.
 [20] Boore J L. Animal mitochondrial genomes [J]. Nucleic Acids Research, 1999, 27(8): 1767-1780.
 [21] Kononov A, Ustyantsev K, Wang B, et al. Genetic diversity among eight *Dendrolimus* species in Eurasia (Lepidoptera: Lasiocampidae) inferred from mitochondrial COI and COII, and nuclear ITS2 markers[J]. BMC Genetics, 2016, 17 (3 Supplement).
 [22] 孙付存,孔祥波,张苏芳,等. 思茅松毛虫 3 个地理种群性信息素成分差异分析[J]. 林业科学研究,2017, 30 (6): 993-998.
 [23] Kong X B, Zhang Z, et al. Female sex pheromone of the Yunnan pine caterpillar moth *Dendrolimus houi*: first (E,Z)-Isomers in pheromone components of *Dendrolimus* spp[J]. Journal of Chemi-

cal Ecology, 2007, 33(7):1316–1327.

学研究进展[J]. 科学通报, 2013, 58(30): 3017–3029.

[24] 吴榴宇. 竹纹枯叶蛾和青黄枯叶蛾线粒体基因组特征及系统发育研究[D]. 南昌: 南昌大学, 2017.

[26] Kong W, Yang J. The complete mitochondrial genome of *Rondotia menciata* (Lepidoptera: Bombycidae) [J]. Journal of Insect Science, 2015, 15(1): 48–48.

[25] 王 维, 孟智启, 石放雄, 等. 鳞翅目昆虫比较线粒体基因组

Sequencing and Analysis of Complete Mitochondrial Genome of *Dendrolimus houi* Lajonquiere (Lepidoptera: Lasiocampidae)

WANG Yao^{1,2}, KONG Xiang-bo², ZHANG Su-fang², LIU Fu², ZHANG Zhen², YAN Shan-chun¹

(1. School of Forestry, Northeast Forestry University, Harbin 150040, Heilongjiang, China;

2. Key Laboratory of Forest Protection of National Forestry and Grassland Administration, Research Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: [Objective] To identify and analyze the mitochondrial genome of *Dendrolimus houi* Lajonquiere, and to study the higher-level phylogenetic relationship of Lepidoptera moths based on the mitochondrial genome.

[Method] The Illumina Hiseq technique was chosen to sequence the purified mitochondrial DNA of *D. houi*. The genes were located and annotated based on the known complete mitochondrial genomes of Lepidopteran species. The tRNA Scan-SE 2.0 was used to predict the secondary structure of the tRNA genes online. Based on the PCGs of the whole nucleotide sequences of mitochondrial genome, the phylogenetic trees of 32 species from 12 families of Lepidoptera moths and the phylogenetic trees of the relative species of the *Dendrolimus* were reconstructed. [Result] The results indicated that the complete mitochondrial genome of *D. houi* is 15 443 bp in length, including 13 protein-coding genes (PCGs), 22 tRNA genes, 2 rRNA genes and a 321 bp A + T-rich region, without gene rearrangement, and showed a higher A + T content (80.0%). All these PCGs except for ND2 and COX1, used ATN as the start codon. Nine PCGs shared the same stop codon TAA, and the stop codons of the other four genes were all defective. COX1, COX2, and ND4 used T as the stop codon and ND3 used TA as the stop codon. Among the 22 tRNA genes, tRNASer (AGN) lacked the DHU arm and could not form a clover structure, while the remaining tRNAs were typical clover structures. The mitochondrial structure was consistent with other insect's mitochondrial genome structures currently available in Lepidoptera. [Conclusion] The phylogenetic analysis indicates that *D. houi* and *D. kikuchii* are completely different relatives, and the relationship between *D. houi* and other *Dendrolimus* species is also far away. The phylogenetic relationship is: (((((*D. tabulaeformis* + *D. punctatus wenshanensis*) + *D. spectabilis* + *D. superans*) + (*D. houi* + *D. kikuchii*)) + *Bombyx mori*). The phylogenetic relationship between the families of Lepidoptera is: ((((((((Lymantriidae + Arctiidae) + Noctuidae) + Notodontidae) + (Geometriidae + (Bombycidae + (Sphingidae + Saturniidae)))))) + Lasiocampidae) + (Pyrallidae + Cramnidae)) + Tortricidae) + Hepialidae).

Keywords: Lepidoptera; Lasiocampidae; *Dendrolimus houi*; mitochondrial genome; phylogeny

(责任编辑:崔 贝)