中国科学 C辑:生命科学 2008 年 第 38 卷 第 8 期: 751~759
www.scichina.com life.scichina.com
《中国科学》杂志社
SCIENCE IN CHINA PRESS

中国野桑蚕线粒体基因组特征

潘敏慧[①], 余泉友[①], 夏玉玲[①], 代方银[①], 刘彦群[①②], 鲁成[①*], 张泽[①], 向仲怀[①]

① 西南大学农业部蚕桑学重点实验室, 重庆 400716;
② 沈阳农业大学生物科学与生物技术学院, 沈阳 110161
* 联系人, E-mail: lucheng@swu.edu.cn

收稿日期: 2008-05-08; 接受日期: 2008-05-26
国家重点基础研究发展计划(批准号: 2005cb121000)和国家自然科学基金(批准号: 30771630)资助项目

摘要 中国野桑蚕(Ch*Bm*)线粒体全基因组序列为 15688 bp, 由 13 个蛋白编码基因、2 个 rRNA 基因、24 个 tRNA 基因和 AT 富集区构成, 比家蚕和日本野桑蚕(Ja*Bm*)线粒体基因组多 2 个 tRNA 基因. 除 *COI* 基因外(含有一个 TTAG 假想起始密码子), 所有蛋白编码基因含有典型的 ATN 起始密码子; 13 个蛋白编码基因中, 有 11 个使用 TAA 作为终止密码子. 除 $tRNA^{Ser(TGA)}$-like 具有四茎环结构外, 其他的 tRNA 具有典型的三叶草结构. Ch*Bm* 的 AT 富集区为 484 bp, 短于 Ja*Bm* 和家蚕. 系统发育分析表明, Ch*Bm* 与家蚕的亲缘关系较近, 与 Ja*Bm* 的亲缘关系较远. Ch*Bm* 与家蚕的分化时间约在(1.08 ± 0.18) ~ (1.41 ± 0.24) 百万年前, 家蚕和 Ja*Bm* 的分化时间约在(1.53 ± 0.20) ~ (2.01 ± 0.26)百万年前, Ch*Bm* 和 Ja*Bm* 的分化时间约在(1.11 ± 0.16) ~ (1.45 ± 0.21)百万年前.

关键词
中国野桑蚕
(Chinese *Bombyx mandarina*, Ch*Bm*)
mtDNA
系统发育
家蚕起源

后生动物线粒体DNA(mtDNA)为环形双链DNA, 基因组大小为 14 ~ 39 kb, 编码 13 个蛋白编码基因、2 个rRNA基因和 22 个tRNA基因[1]. 尽管在后生动物线粒体基因组中基因数目高度保守, 但也有一些例外的报道, 例如双翅目的*Chrysomya putoria*和膜翅目的红光熊蜂(*Bombus ignites*)就具有不同数目的tRNA基因[2,3]. 昆虫线粒体基因组AT富集区, 它含有一个在序列和长度上比其他区域更高度可变[4~6].

家蚕(*Bombyx mori*)是唯一的驯化昆虫, 它是由5000 ~ 10000 年前的古野桑蚕驯化而来, 这些野桑蚕广泛分布在亚洲各地[7]. 然而在野桑蚕中存在两种染色体数目不同的野桑蚕, 一种为日本野桑蚕(Japanese *Bombyx mandarina*, Ja*Bm*), 染色体为 27 条, 生活在日本和韩国南部部分地区[8,9]; 另一种为中国野桑蚕(Chinese *Bombyx mandarina*, Ch*Bm*), 与驯化的家蚕一样, 具有28条染色体[10,11]. 尽管一些研究表明家蚕与野桑蚕的亲缘关系很近[12,13], 但家蚕起源于何种野桑蚕还不清楚.

mtDNA 广泛应用在系统发育和群体遗传研究上[14], 因为mtDNA属于母性遗传, 并比核基因组具有更高的突变率[15,16]. 目前, Yukuhiro等人[12]已进行了Ja*Bm*的线粒体基因组测序, 发现Ja*Bm*与家蚕的线粒体基因组存在显著的差异, 参照甲虫*nad5* 基因碱基突变率, 估算出Ja*Bm*与家蚕大约在 7.1 百万年前分开. 于此, 我们公布了Ch*Bm*的线粒体基因组序列, 并与Ja*Bm*和家蚕的线粒体基因组序列进行了比较. 利用家蚕、野桑蚕和柞蚕(*Antheraea pernyi*)的线粒体基因组的蛋白质编码基因序列构建的系统发育树表明, 含有相同数目(28 条)染色体的家蚕和中国野蚕具有更近的亲缘关系, 而家蚕与含有 27 条染色体的Ja*Bm*的亲缘关系相对较远.

1 材料与方法

1.1 昆虫和 mtDNA 的准备

Ch*Bm*从陕西省安康地区收集. 野桑蚕mtDNA从野桑蚕的蛹中抽提, 制备方法参照《分子克隆实验指南》第 2 版进行[17].

1.2 引物设计、PCR 和测序

利用 Oligo (version 6.0)软件参照 Ja*Bm* 全基因组序列进行 PCR 引物设计. PCR 采用下列方法进行: 94℃预变性 2 min; 94℃变性 30 s, 34 ~ 50℃退火 40 s, 72℃延伸 45 ~ 60 s, 重复 35 个循环; 最后 72℃延伸 10 min. PCR 产物用回收试剂盒(QIAquick Gel Extraction Kit, Germany) 纯化回收后直接在 ABI PRISM^{TM} 310 Genetic Analyzer (PE Applied Biosystems, USA)上测序.

1.3 序列分析

利用CLUSTAL X[18]软件对Ch*Bm*和其他昆虫线粒体全基因组进行了序列比对, 发现Ch*Bm*线粒体基因组中同样存在线粒体基因和AT富集区域. 线粒体蛋白编码基因使用无脊椎动物线粒体密码子, tRNA基因与其他昆虫(包括家蚕、Ja*Bm*和红光熊蜂等昆虫)线粒体tRNA基因比对后进行确定[3,12]并用其反密码子序列进行确认. tRNA基因的二级结构通过tRNAscan-SE 软件[19]进行预测. 蛋白编码基因和rRNA基因通过与公布的家蚕线粒体相应基因序列的相似性进行确定. Ch*Bm*线粒体基因组序列已提交到GeneBank, 登录号为AY301620.

1.4 系统发育分析

利用完整的家蚕(C108: AB070264, Aojuku: AB083339, 夏芳: AY048187 和Backokjam: AF149768)、Ja*Bm* (NC_003395)和柞蚕(NC_004622)等昆虫的线粒体基因组与Ch*Bm*进行比较, 利用线粒体基因组的 13 个编码蛋白基因进行系统发育分析. 7 个昆虫的线粒体基因组序列用Clustal X进行比对, 利用MEGA version 2.1[20]构建系统发育树, 利用邻接(Neighbor- joining, NJ)法进行遗传距离修正. 利用TREE-PUZZLE 5.0[21]构建最大似然(Maximum Likelihood, ML)树.

7 个昆虫的*COI* 和 *COII*基因分别连接后用Clustal X 1.8 软件进行比对, 利用核酸替代的Kimura双参数模型[22]和MEGA version 2.1 软件进行成对遗传距离估算. 以凤蝶 (*Papilionidae Papilio*)*COI*+*COII*基因的碱基替代频率 7.8×10^{-9} ~ 10.2×10^{-9}/年作为分子钟进行家蚕和其他蚕蛾科昆虫间的分化时间估算. 估算的分化时间(T)与分子钟(r)和遗传距离(K)之间的关系为: $T = K / (2r)$[23].

2 结果

2.1 Ch*Bm* mtDNA 基因组结构和组成

完整的Ch*Bm* mtDNA测序结果提交到GeneBank(登录号: AY301620). Ch*Bm* mtDNA全长15688 bp, 包括 13 个蛋白编码基因、2 个rRNA基因、24 个tRNA基因和一段富含AT的非编码区域. 除多出的 2 个tRNA基因(tRNA$^{Ser(UGA)}$-like 和 tRNA$^{Ile(UAU)}$-like)外, Ch*Bm*与家蚕和Ja*Bm*线粒体基因组有着相同的基因排列顺序[12]. 除多出的 2 个tRNA基因序列外, Ch*Bm* mtDNA基因间在 11 个位点有 45 bp的重叠, 基因间有 875 bp的内含子序列, 包括 17 个分隔区域(391 bp)和AT富集区域(484 bp). 内含子的大小从 1 ~ 59 bp不等, 最大的内含子位于tRNA^{His} 和 ND4 基因之间(表 1).

Ch*Bm* mtDNA 组成偏好于碱基 A 和 T, AT 含量高达 81.59%. Ja*Bm* 和家蚕具有相似的碱基偏好性, AT 含量分别为 81.68%和 81.32%(表 2). Ch*Bm* mtDNA AT 富集区域的 AT 含量为 94.42%, 低于 Ja*Bm* 的 95.18%和家蚕的 95.39%. Ch*Bm*, Ja*Bm* 和家蚕线粒体基因组的 AT 和 GC 组成偏好性(content skews)分别呈现出弱阳性和弱阴性, 表明它们偏好使用碱基 A 和 C, 但原因还不清楚.

2.2 蛋白编码基因

Ch*Bm* mtDNA的 13 个蛋白编码基因使用典型的ATN启动子, 其中 6 个使用ATA, 5 个使用ATG, *ND5*使用ATT. 然而, *COI*基因可能与Ja*Bm*的一样, 使用TTAG 4 碱基启动子[12].

Ch*Bm* mtDNA 的 13 个蛋白编码基因与 Ja*Bm* 和家蚕的具有相似的特征. 11 个蛋白编码基因具有 TAA 终止密码子, 但 *COI* 和 *COII* 基因使用 1 个不完整

表 1 Ch*Bm*, Ja*Bm* 和家蚕 C108 的线粒体基因组结构

基因或区域	位置	大小/bp	起始密码子	终止密码子	3 端间隔 [a)]	编码链 [b)]	Ja*Bm*/bp	*B. mori* C108/bp
COIII	(15134 ~ 15682, 1 ~ 240)	789	ATG	TAA	2	H	798	798
$tRNA^{Gly}$	243 ~ 311	69			3	H	66	66
ND3	315 ~ 665	351	ATA	TAA	53	H	351	351
$tRNA^{Ala}$	719 ~ 786	68			50	H	64	68
$tRNA^{Arg}$	837 ~ 900	64			0	H	64	64
$tRNA^{Asn}$	901 ~ 967	67			0	H	67	65
$tRNA^{Ser}$	968 ~ 1034	67			10	H	67	67
$tRNA^{Glu}$	1045 ~ 1107	63			–1	H	64	65
$tRNA^{Phe}$	1107 ~ 1173	67			4	L	67	67
ND5	1178 ~ 2896	1719	ATT	TAA	16	L	1719	1719
$tRNA^{His}$	2913 ~ 2981	69			59	L	67	67
ND4	3041 ~ 4384	1344	ATG	TAA	–1	L	1341	1341
ND4L	4384 ~ 4674	291	ATG	TAA	4	L	291	291
$tRNA^{Thr}$	4679 ~ 4744	66			–1	H	66	65
$tRNA^{Pro}$	4744 ~ 4812	69			1	L	68	66
ND6	4814 ~ 5344	531	ATA	TAA	55	H	531	531
cytb	5400 ~ 6557	1158	ATA	TAA	–1	H	1158	1152
$tRNA^{Ser}$	6557 ~ 6622	66			25	H	66	66
ND1	6648 ~ 7592	945	ATA	TAA	–6	L	945	945
$tRNA^{Leu}$	7587 ~ 7657	71			33	L	71	70
16S rRNA	7691 ~ 9040	1350			0	L	1377	1378
$tRNA^{Val}$	9041 ~ 9111	71			0	L	72	69
12S rRNA	9112 ~ 9895	784			0	L	783	783
AT rich region	9896 ~ 10379	484			0	H	747	494
$tRNA^{Ser(TGA)}$-like	9965 ~ 10055	91	–	–	–	H	–	–
$tRNA^{Met}$	10380 ~ 10447	68			–2	H	68	68
$tRNA^{Ile}$	10446 ~ 10511	66			–3	H	66	66
$tRNA^{Gln}$	10509 ~ 10577	69			48	L	69	69
$tRNA^{Ile(TAT)}$-like	10570 ~ 10639	70	–	–	–14	H	–	–
ND2	10626 ~ 11648	1023	ATA	TAA	5	H	1023	1023
$tRNA^{Trp}$	11654 ~ 11723	70			–8	H	70	70
$tRNA^{Cys}$	11716 ~ 11782	67			6	L	67	67
$tRNA^{Tyr}$	11789 ~ 11854	66			17	L	66	66
COI	11872 ~ 13402	1531	TTAG	T++	0	H	1531	1531
$tRNA^{Leu}$	13403 ~ 13469	67			0	L	67	67
COII	13470 ~ 14151	682	ATG	T++	0	H	682	682
$tRNA^{Lys}$	14152 ~ 14222	71			–1	H	71	71
$tRNA^{Asp}$	14222 ~ 14288	67			0	H	67	67
ATP8	14289 ~ 14450	162	ATA	TAA	–7	H	162	162
ATP6	14444 ~ 15121	678	ATG	TAA	–14	H	678	678

a) 负数表示基因间的碱基重叠数; b) H 表示线粒体重链, L 表示线粒体轻链

表 2　Ch*Bm*, Ja*Bm* 和 *C108* 线粒体不同基因中碱基组成的偏好性

物种	全线粒体基因组大小/ bp	A/%	G/%	T/%	C/%	(AT)%	AT 偏好性	GC 偏好性
B. mori C108	15643	43.05	7.32	38.27	11.36	81.32	0.058	–0.215
Ja*Bm*	15928	43.08	7.21	38.60	11.11	81.68	0.055	–0.213
Ch*Bm*	15682	43.11	7.40	38.48	11.01	81.59	0.057	–0.196
PCGs								
B. mori C108	11170	42.91	8.17	36.66	12.27	79.57	0.079	–0.201
Ja*Bm*	11194	42.78	8.14	36.86	12.22	79.64	0.074	–0.200
Ch*Bm*	11197	42.83	8.26	37.04	11.87	79.87	0.072	–0.179
tRNAs								
B. mori C108	1462	42.20	7.80	39.33	10.67	81.53	0.035	–0.155
Ja*Bm*	1463	41.90	7.79	39.71	10.59	81.61	0.026	–0.152
Ch*Bm*	1472	41.78	7.81	39.95	10.46	81.73	0.022	–0.145
rRNA								
B. mori C108	2158	43.74	4.59	41.06	10.61	84.84	0.032	–0.397
Ja*Bm*	2160	43.89	4.63	41.30	10.19	85.19	0.030	–0.375
Ch*Bm*	2134	43.86	4.78	41.50	10.31	84.91	0.028	–0.366
A+T rich region								
B. mori C108	499	44.69	1.60	50.70	3.01	95.39	–0.063	–0.305
Ja*Bm*	747	45.52	2.41	49.67	2.41	95.18	–0.043	0.000
Ch*Bm*	484	46.49	2.69	47.93	2.89	94.42	–0.015	–0.036

终止密码T. Ch*Bm*线粒体基因与Ja*Bm*一样, 比家蚕的*cytb*基因长 6 bp[12]. Ch*Bm*线粒体基因*ND4* 比Ja*Bm*和家蚕的*ND4* 基因长 3 bp. 其他 11 个蛋白编码基因长度相同(表 1).

在 Ch*Bm*, Ja*Bm* 和家蚕的 13 个线粒体蛋白编码基因中, AT 组成偏好性呈弱阳性, GC 组成偏好性呈弱阴性(表 2). 尽管在 3 个线粒体基因组中的三联体密码子过多使用 A 和 T, 但密码子每一位点有不同的 AT/GC 使用偏好性. 在密码子第一位, Ch*Bm*, Ja*Bm* 和家蚕分别偏好使用 A 和 G, A 和 C, T 和 C. 在密码子第二位, 都偏好使用 T 和 G. 在密码子第三位, Ch*Bm* 和家蚕偏好使用 T 和 G, 而 Ja*Bm* 偏好使用 T 和 C (表 3).

2.3　tRNA 和 rRNA

Ch*Bm* 线粒体 tRNA 基因通过 tRNAscan-SE Search Server 软件预测. Ch*Bm* 和 Ja*Bm* 线粒体基因组的 $tRNA^{Gly}$, $tRNA^{Val}$, $tRNA^{Ala}$, $tRNA^{His}$ 和 $tRNA^{Glu}$ 基因存在差异, 并预测了两种野桑蚕的 $tRNA^{Gly}$, $tRNA^{Leu(TAG)}$, $tRNA^{Pro}$, $tRNA^{Val}$, $tRNA^{His}$ 和 $tRNA^{Glu}$ 基因结构上的差异(图 1), 这些差异主要是由碱基的插入和缺失引起.

Ch*Bm* 比 Ja*Bm* 和家蚕线粒体基因组多出 2 个 tRNA-like 基因(表 1), $tRNA^{Ser(TGA)}$-like 基因由 H 链编码, 位于 AT 富集区域 (9965 ~ 10055 bp), 形成 3 个小环和 1 个大环的四茎环结构; $tRNA^{Ile(TAT)}$-like 基因位于

表 3　Ch*Bm*, Ja*Bm* 和家蚕 C108 线粒体蛋白质编码基因中密码子的使用情况

物种	第一位密码子				第二位密码子				第三位密码子			
	A	T	G	C	A	T	G	C	A	T	G	C
Ch*Bm*	38.2	36.9	16.1	8.8	22.9	50.7	13.4	13.0	40.6	52.7	3.4	3.3
Ja*Bm*	37.9	36.8	10.9	14.4	25.3	47.9	16.1	10.7	41.5	47.7	4.2	6.6
*B.mor*i C108	37.3	37.6	9.4	15.7	22.2	48.6	15.9	13.3	43.6	49.2	4.5	2.7

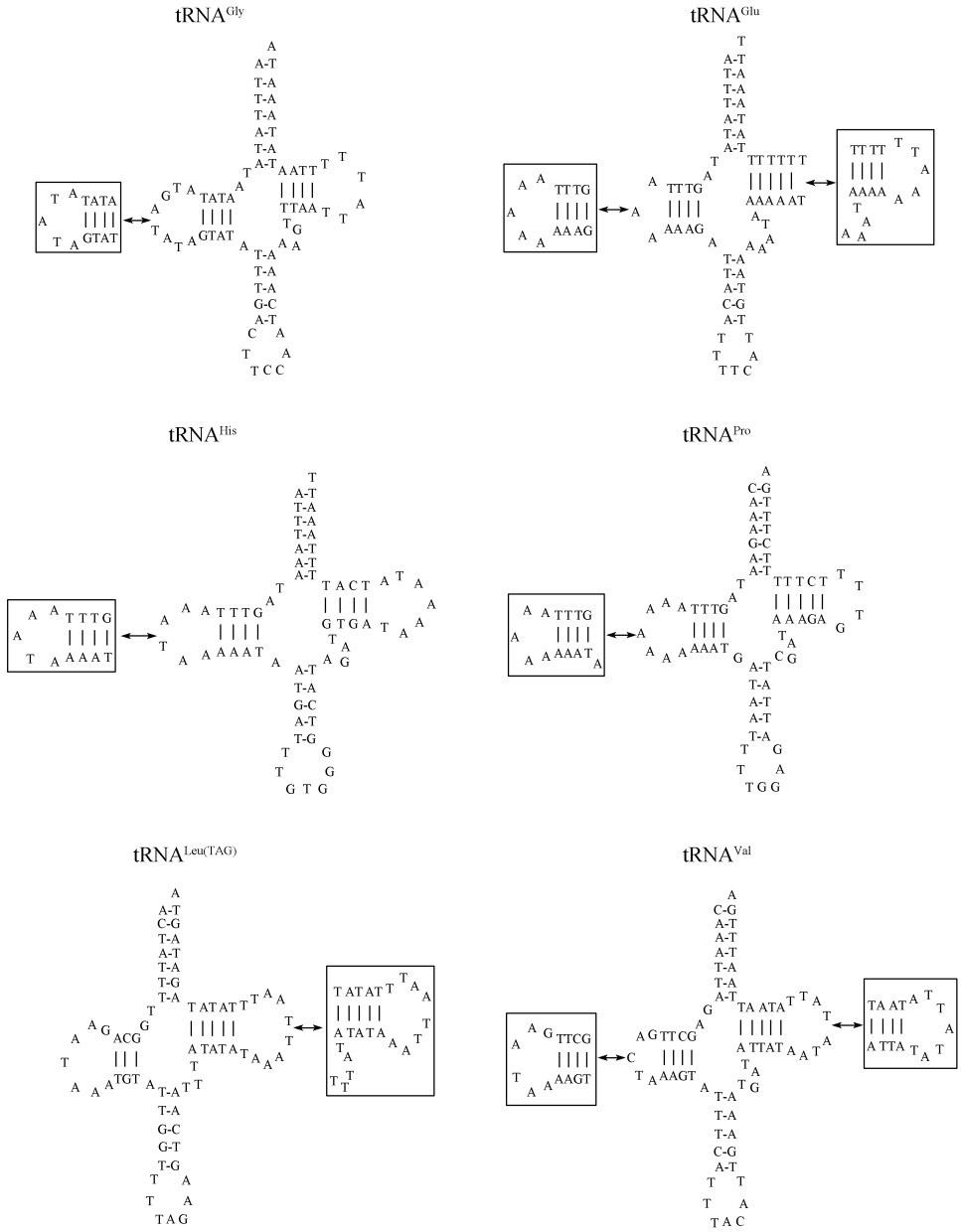

图 1 Ch*Bm* 和 Ja*Bm* 线粒体 tRNA 二级结构的差异; Ch*Bm* tRNA 的三叶草结构在图中显示出来, 框中部分为 Ja*Bm* tRNA 的差异部分

10570 ~ 10639 bp之间, 分别与tRNA^{Gln}和*ND2* 基因重叠 9 bp和 13 bp, 形成典型的三叶草结构(图 2). 尽管后生动物线粒体基因组的基因数目高度保守, 但还是有一些例外, 如双翅目的*C. putoria*就具有 1 个额外的*tRNA^{Ile}* 基因[2]. 另外, 其他的膜翅目昆虫也具有一些数目不等的tRNA基因[3,24,25]. 然而, 这是首次在鳞翅目昆虫发现*tRNA*基因数目发生变化.

Ch*Bm* mtDNA 的 2 个 rRNA 基因位于 tRNA^{Leu} 和 AT 富集区域之间. Ch*Bm* 16S rRNA 基因大小为 1350 bp, 小于 Ja*Bm* 的 1377 bp 和家蚕的 1378 bp. 而 Ch*Bm* 的 12S rRNA 基因大小为 784 bp, 与 Ja*Bm* 的 12S rRNA 基因(783 bp) 和家蚕的 12S rRNA 基因

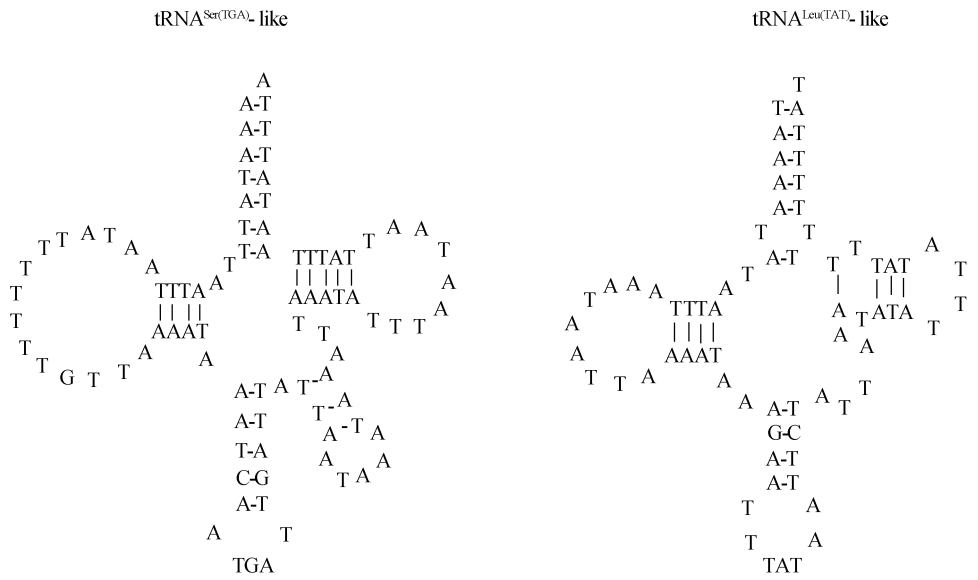

图 2　Ch*Bm* 中 tRNA-like 基因的预测结构

(783 bp)大小相似 (表 1). Ch*Bm*, Ja*Bm* 和家蚕的 tRNA 基因与 rRNA 基因都偏向使用碱基 A 和 C(表 2).

2.4　AT 富集区域

不同昆虫的AT富集区域大小差异较大, 即使在同一个种内也是如此. 亲缘关系相近物种AT富集区域大小的差异主要由AT-rich区域中的串联重复数目不同造成[26]. Ch*Bm*和Ja*Bm*线粒体明显的差异出现在二者AT富集长度的不同(表 2), 前者AT富集区域大小为 484 bp, 而后者则为 747 bp. Ch*Bm* AT富集区域大小与家蚕(494 ~ 499 bp)接近[12,27]. Ja*Bm* 的AT富集区域含有 3 个 126 bp的串联重复, 但在Ch*Bm*和家蚕的AT富集区域中仅发现 1 个串联重复[12,13].

2.5　系统发育分析

从NCBI下载不同品种的家蚕、Ja*Bm*和柞蚕的线粒体全基因组用于系统发育分析. 将 13 个线粒体蛋白编码基因推测氨基酸序列连接后用Clustal X进行比对, 获得 1 个有 3748 个氨基酸的对齐序列, 利用这个对齐序列进行NJ和ML计算, 获得了最高支持度的ML树(图 3). 四个家蚕品种形成一簇, 然后再分别与Ch*Bm*, Ja*Bm*和柞蚕聚在一起, 表明Ch*Bm*与家蚕的亲缘关系最近, 家蚕与Ch*Bm*在驯化之前与Ja*Bm*分化开. 这个结果不同于核基因组编码的*apolipophorin*-III基因和线粒体*cyt b*基因的研究结果[12,28].

家蚕品种间、Ch*Bm*-家蚕间、Ja*Bm*-家蚕间和Ch*Bm*-Ja*Bm*间的平均遗传距离分别为 0.0007, 0.0227, 0.0319 和 0.0225(表 4). Ch*Bm*-家蚕间和 Ch*Bm*-Ja*Bm*间的遗传距离相近, Ja*Bm*-家蚕之间的遗传距离稍微大于 Ch*Bm*-Ja*Bm* 间的遗传距离.

凤蝶线粒体*COI* + *COII*基因的碱基替代频率为 7.8×10^{-9} ~ 10.2×10^{-9}/年[29]. 利用此碱基替代频率作为分子钟, 根据表 4 的碱基替代率估算了家蚕和其近缘种的分化时间(表 5). 家蚕-Ch*Bm* 和家蚕-Ja*Bm*的分化时间分别大约在 (1.08 ± 0.18) ~ (1.41 ± 0.24)百万年和(1.53 ± 0.20) ~ (2.01 ± 0.26)百万年前; Ch*Bm* 和 Ja*Bm* 分化时间大约在(1.11 ± 0.16) ~ (1.45 ± 0.21)百万年前, 并且二者与柞蚕的分化时间大致相同.

3　讨论

目前在亚洲分布着两种野桑蚕, 具有 27 条染色体的Ja*Bm*分布在日本和韩国南部临海的部分地区, 另外一种具有 28 条染色体的野桑蚕分布在中国和亚洲的其他区域. 地质调查和这两种野桑蚕早期的核糖体RNA研究结果表明, 二者的分化最早出现在朝鲜海峡的北部[30~32]. 近期的家蚕与Ja*Bm*杂交结果

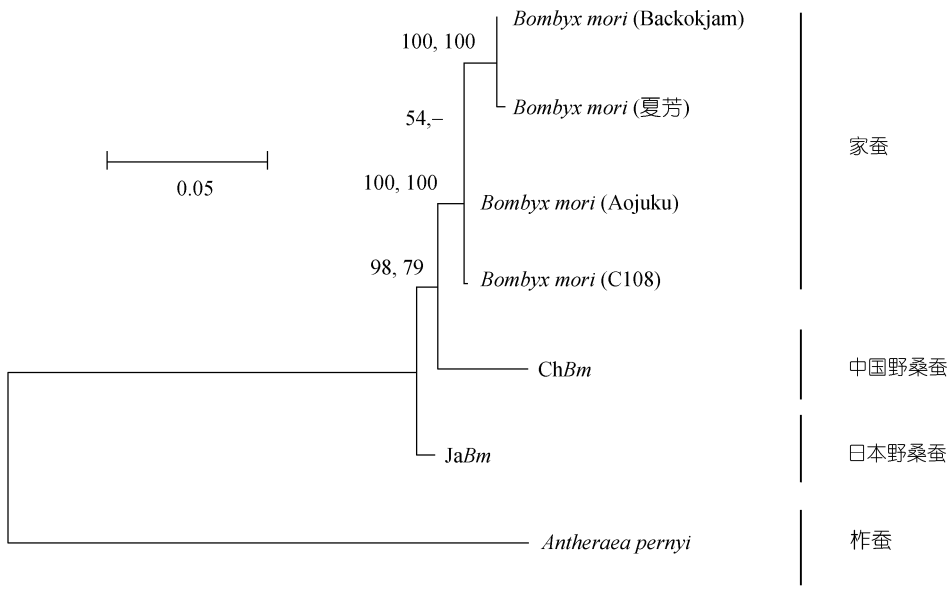

图 3 利用氨基酸对齐序列构建的最大似然树

ML 和 NJ 法获得了相似的拓扑结构, 节点上的数字代表 ML 和 NJ 法分析的支持度, "–"表示在 ML 树中的支持度小于 50%

表 4 利用线粒体 *COI* + *COII* 基因估算的遗传距离和遗传距离的标准误差 a)

物种	*B. mori* (Aojuku)	*B. mori* (Bachokjam)	*B. mori* (C108)	*B. mori* (夏芳)	Ch*Bm*	Ja*Bm*	柞蚕
B. mori (Aojuku)		0.0000	0.0005	0.0006	0.0037	0.0041	0.0087
B. mori (Bachokjam)	0.0000		0.0005	0.0006	0.0037	0.0041	0.0087
B. mori (C108)	0.0005	0.0005		0.0008	0.0037	0.0041	0.0087
B. mori (夏芳)	0.0009	0.0009	0.0014		0.0038	0.0041	0.0087
Ch*Bm*	0.0225	0.0225	0.0220	0.0235		0.0033	0.0088
Ja*Bm*	0.0316	0.0316	0.0312	0.0326	0.0225		0.0086
Anpe	0.1337	0.1337	0.1332	0.1348	0.1310	0.1310	

a) 对角线上为遗传距离; 对角线下为遗传距离的标准误差

表 5 四种产丝昆虫的进化分化时间(单位: 百万年)

物种	*B. mori* (C108)	Ch*Bm*	Ja*Bm*
Ch*Bm*	(1.08 ± 0.18) ~ (1.41 ± 0.24)		
Ja*Bm*	(1.53 ± 0.20) ~ (2.01 ± 0.26)	(1.11 ± 0.16) ~ (1.45 ± 0.21)	
柞蚕	(6.53 ± 0.43) ~ (8.55 ± 0.56)	(6.43 ± 0.43) ~ (8.40 ± 0.56)	(6.43 ± 0.42) ~ (8.41 ± 0.55)

显示出 Ja*Bm* 的 M-染色体能与家蚕的 2 条染色体反应[9]. 由于两种野桑蚕很难从形态上进行区分, 一直使用相同的学名*B. mandarina*. Ch*Bm* 线粒体全基因组表明它与家蚕的亲缘关系更近(图 3). 另外, Ch*Bm* AT富集区域的长度与家蚕相近, 而与Ja*Bm*的相差很大. 根据这些差异, 我们认为应对具有 27 条染色体的Ja*Bm*进行重新分类, AT富集区域的长度和染色体数目的差异应是二者特征差异.

尽管已有人利用核基因和mtDNA序列对家蚕进行了系统发育研究[12,28], 但这些研究并未就家蚕的起源得出共同的结论. 因此, 家蚕是单起源还是多起源还不能确定. Yukuhiro等人[12]认为Ja*Bm*可能是家蚕的祖先, 但Ch*Bm* 和 Ja*Bm*线粒体基因组比较结果和系统发育分析结果表明, 家蚕应是Ch*Bm* 的直系后代而不是 Ja*Bm* 的后代(图 3), 这与家蚕的驯化历史一致[7,33].

Yukuhiro 等人[12]采用大步甲虫(*Coleoptera Car-*

abidae)的*nad5* 基因作为分子钟估算出家蚕C108 与Ja*Bm*的 分化时间大约在 7.1 百万年前. 然而, 甲虫与蚕蛾科的亲缘关系很远, 线粒体基因组在不同属间的碱基替代频率变化就很大[29], 因此用甲虫的*nad5* 基因作为分子钟估算家蚕C108 与Ja*Bm*的分化时间不很恰当. 而本文采用与家蚕同属于鳞翅目的凤蝶线粒体*CO* Ⅰ和*CO* Ⅱ基因的碱基替代频率7.8×10^{-9} ~ 10.2×10^{-9}为分子钟[29], 估算了家蚕C108与Ch*Bm*, Ch*Bm* 和Ja*Bm*之间的分化时间(表 5). Ch*Bm* 和家蚕C108 的分化时间大约在 (1.08 ± 0.18) ~ (1.41 ± 0.24)百万年前, Ja*Bm*和 家蚕C108 的分化时间大约在(1.53 ± 0.20) ~ (2.01 ± 0.26)百万年前. 这些估算的分化时间小于Yukuhiro等人[12]估算出的分化时间, 这是由于他们使用了一个不恰当的分子钟高估了Ja*Bm*和家蚕C108 的分化时间. 然而历史记录表明家蚕是在中国五千年前驯化的[7,33], 这个时间远远短于 1.08 百万年, 因此, Ch*Bm*与家蚕线粒体基因发生的显著差异可能与长期驯化过程中的强烈人工选择有关.

本研究估算出家蚕与 Ja*Bm* 的分化时间大约在(1.53 ± 0.20) ~ (2.01 ± 0.26)百万年前, 这可能是一个比较合理的估算时间. 大约在两百万年前, 日本群岛属于冰河期, 黄海和韩国-对马岛海峡变成了日本群岛西南部与大陆相连的陆地[34]. 在那时, 野桑蚕可能从朝鲜进入日本 (1.5 ~ 2.1 百万年前), 然后被隔离在这个区域, 进化成含有 27 条染色体的Ja*Bm*. 在冰河时代的后期(0.15 ~ 0.01 百万年前), 最多在二万年前, 日本群岛的西南部与朝鲜经过一个狭窄的陆地桥相连[35,36]. 那时一些Ja*Bm*可能从日本群岛迁移到朝鲜半岛, 并在与日本群岛环境相似的地区存活下来. 同时, 朝鲜半岛也可能有野桑蚕迁移到日本群岛, 但却没有在日本存活下来, 可能是这些野桑蚕不能适应日本群岛环境的缘故, 但这需要更多的证据. 这些家蚕与其祖先的比较研究结果可为家蚕的遗传进化研究提供有用的信息.

参考文献

1 Wolstenholme D R. Animal mitochondrial DNA: structure and evolution. Int Rev Cyt, 1992, 141: 173—216[DOI]

2 Junqueira A C. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). Gene, 2004, 339: 7—15[DOI]

3 Cha S Y, Yoon H J, Lee E M, et al. The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignites* (Hymenoptera: Apidae). Gene, 2007, 392: 206—220[DOI]

4 Fauron C M R, Wolstenholme D R. Extensive diversity among *Drosophila* species with respective to nucleotide sequences within the adenine + thymine rich region of mitochondrial DNA molecules. Nucleic Acids Res, 1980, 11: 2439—2453[DOI]

5 Lewis D L, Farr C L, Kaguni L S. *Drosophila melanogaster* mitochondrial DNA, completion of the nucleotide sequence and evolutionary comparisons. Insect Mol Biol, 1995, 4: 263—278[DOI]

6 Renfu S, Nick J H, Campbell H, et al. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol Biol Evol, 2001, 18: 858—865

7 Goldsmith M R, Shimada T, Abe H. The genetics and genomics of the silkworm, *Bombyx mori*. Annu Rev Entomol, 2005, 50: 71—100[DOI]

8 Sasaki C. On the affinity of our wild and domestic silkworms. Ann Zool Jpn 1898, 2: 33—41

9 Banno Y, Nakamura T, Nagashima E, et al. M chromosome of the wild silkworm, *Bombyx mandarina* (*n*=27), corresponds to two chromosomes in the domestication silkworm, *Bombyx mori* (*n*=28). Genome, 2004, 47: 96—101[DOI]

10 Kawaguchi E. Zytologiche untersuchungen am seidenspinner und seinen verwandten. I. Gametogenes von *Bombyx mori* L. und *B. mandarina* M. und ihrer Bestarde. Z Zellforsch Mikrosk Anat, 1928, 7: 154—156[DOI]

11 Astaurov B L, Golisheva M D, Roginskaya I S. Chromosome complex of Ussuri geographical race of *Bombyx mandarina* M. with spacial reference to the problem of the origin of the domesticated silkworm, *Bombyx mori* L. Cytology, 1959, 1: 327—332 [in Russian]

12 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. Mol Biol Evol, 2002, 19: 1385—1389
13 Arunkumar K P, Metta M, Nagaraju J. Molecular phylogeny of silkworm reveals the origin of domesticated silkmoth, *Bombyx mori* from Chinese *Bombyx mandarina* and paternal inheritiance of *Antheraea proylei* mitochondrial DNA. Mol Phylogenet Evol, 2006, 40: 419—427[DOI]
14 Avise J C. Phylogeography: The History and Formation of Species. Cambridge：Harvard University Press, 2000. 20—147
15 Moritz C, Dowling T E, Brown W M. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Ann Rev Ecol Syst, 1987, 18: 269—292[DOI]
16 Pesole G, Gissi C, Saccone C. Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol, 1999, 48: 427—434[DOI]
17 Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning: A Laboratory Manual. 3th ed. New York: Cold Spring Harbor Lab Press, 2001. 138—153
18 Thomson J D, Gibson T J, Plewniak F, et al. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res, 1997, 24: 173—216
19 Lowe T M, Eddy S R. tRNA-scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res, 1997, 25: 955—964[DOI]
20 Kumar S, Tamura K, Jakobsen I B, et al. MEGA 2: Molecular Evolution Genetics Analysis Software. Bioinformatics, 2001, 17: 1244—1245[DOI]
21 Schmidt H A, Strimmer K, Vingron M, et al. TREE-PUZZLE: Maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics, 2002, 18: 502—504[DOI]
22 Kumar M. A simple method for estimating evolution rate of base substitution through comparative studies of nucleotide sequence. J Mol Evol, 1980, 16: 111—120[DOI]
23 Fu Y X, Li W H. Estimating the age of the common ancestor of a sample of DNA sequences. Mol Biol Evol, 1997, 14: 195—199
24 Shao R, Barker S C. The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol Biol Evol, 2003, 20: 362—370[DOI]
25 Thao M L, Baumann L, Baumann P. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). BMC Evol. Biol, 2004, 4: 25[DOI]
26 Zhang D X, Hewitt G M. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol, 1997, 25: 99—120
27 Lu C, Liu Y Q, Liao S Y, et al. Complete sequence dternination and analysis of *Bombyx mori* mitochondrial genome. Journal of Agricultural Biotechnology, 2002, 10: 163-170 (in Chinese)
28 鲁成, 刘运强, 廖顺尧, 等. 家蚕线粒体基因组全序列测定与分析. 农业生物技术学报, 2002, 10: 163—170
29 Li A, Zhao Q, Tang S, et al. Molecular phylogeny of the domesticated silkworm, *Bombyx mori*, based on the sequences of mitochondrial cytochromeb genes. J Genet, 2005, 84: 137—142[DOI]
30 Zakharov E V, Caterino M S, Sperling F A. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus Papilio (Lepidoptera: Papilionidae). Syst Biol, 2004, 53: 193—215[DOI]
31 Maekawa H, Takada N, Mikitani K, et al. Nucleolus organizers in the wild silkworm *Bombyx mandarina* and domesticated silkworm *B. mori*. Chromosoma, 1988, 96: 263—269[DOI]
32 Nakamura T, Banno Y, Nakada T, et al. Geographic dimorphism of the wild silkworm, *Bombyx mandarina*, in the chromosome number and the occurrence of a retroposon – like insertion in the arylphorin gene. Genome, 1999, 42: 1117—1120[DOI]
33 Shimada T, Kurimoto Y, Kobayashi M. Phylogenetic relationship of silkmoths inferred from sequence data of the arylphorin gene. Mol Phylogenet Evol, 1995, 4: 223—234[DOI]
34 Zhou Y. General Entomology. 2nd ed. Beijing: High Education Publication House, 1958
35 Tominaga O, Su Z H, Kim C G, et al. Formation of the Japanese *Carabina fauna* inferred from a phylogenetic tree of mitochondrial *ND5* gene sequences (Coleoptera, Carabidae). J MolEvol, 2000, 50: 541—549
36 Oshima K. The history of straits around the Japanese Islands in the late-quaternary. Quaternary Res, 1990, 29: 193—208
37 Matsui H, Tada R, Oba T. Low-salinity isolation event in the Japan Sea in response to eustatic sea-level drop during LGM: reconstruction based on salinity-balance model. Quaternary Res, 1998, 37: 221—233