



论文

蚕系统发生及进化历史分析

孙伟^①, 于红松^①, 沈以红^①, Banno Yutaka^②, 向仲怀^①, 张泽^{①③*}

① 西南大学蚕学与系统生物学研究所, 重庆 400716;

② 九州大学遗传资源研究所生物资源及生物环境科学研究院, 福岡 812-8581, 日本;

③ 重庆大学农学及生命科学研究院, 重庆 400044

* 联系人, E-mail: zezhang@swu.edu.cn

收稿日期: 2012-01-08; 接受日期: 2012-05-14

国家自然科学基金(批准号: 30671587)、高等学校学科创新引智计划(批准号: B07045)和西南大学博士创新基金(批准号: Kb2009019)资助项目

摘要 古丝绸之路是连接古亚洲和古欧洲的纽带. 作为重要媒介, 家蚕和丝绸在其中发挥着重要作用. 家蚕是目前唯一一种被完全驯化的昆虫. 然而, 利用分子数据研究家蚕起源及进化历史的研究还很少, 尤其是家蚕及其祖先野桑蚕的分化时间还有待确定. 作为重要的驯化物种, 家蚕的群体动力学变化历程也鲜有报道. 本研究利用线粒体和核基因 DNA 序列推断家蚕及其近缘物种的系统发生关系和进化历史. 所有系统发生分析结果表明, 家蚕和中国野桑蚕具有更近的亲缘关系. 家蚕的驯化时间大约为 4100 年. 中国野桑蚕大约在 23600 年前与日本野桑蚕分开, 这个时间与化石证据和历史记载一致. 另外, 研究结果表明, 家蚕在大约 1000 年前经历了群体扩张. 本文对蚕的分化时间及蚕群体动力学进行研究, 不仅有利于驯化动物的遗传分析和鳞翅目昆虫的系统发生分析, 而且有利于理解人类文明的传播.

关键词家蚕
驯化
分化时间
群体动力学
群体扩张

丝绸产业在许多发展中国家, 如中国、印度和巴西仍然是重要的支撑产业. 家蚕(*Bombyx mori*)是丝绸产业的基础. 现今, 全世界有超过 1000 个家蚕品系. 一般认为, 家蚕由其祖先野桑蚕(*Bombyx mandarina*)驯化而来^[1-4]. 目前, 存在两种野桑蚕: 一种生活在中国及俄罗斯远东地区, 通常称为中国野桑蚕(染色体数目: $2n=56$, 和家蚕的染色体数目一致); 另一种栖息在日本及朝鲜半岛地区, 称为日本野桑蚕(染色体数目: $2n=54$, 比家蚕少一对染色体)^[1,5]. 考古学和细胞遗传学研究表明, 家蚕由中国野桑蚕驯化而来, 分子系统发生研究也支持这些结论^[4,6-9]. 而后, 家蚕及养蚕技术在大约公元前 200 年传至韩国和日本, 随后伴随丝绸之路传至中亚和欧洲^[10]. 在此期间, 丝绸和养蚕技术的传播在古代国家之间的贸易

和文化交流方面起着重要作用. 因此, 家蚕从中国传播至其他国家的历史在一定程度上代表了古亚洲和古欧洲的贸易和文化交流的历史. 尽管现在有很多书籍记载这一历史, 但是分子水平的证据仍然十分缺乏. 最近, Pan 等人^[11]利用凤蝶科(Papilionidae)昆虫线粒体 *COI* 和 *COII* 基因的核苷酸替换速率计算了蚕的分化时间, 表明家蚕的驯化时间大约发生在 1.08~1.41 Ma (million years ago). 另一研究^[12]利用鞘翅目步甲科(Carabidae)昆虫的线粒体 *Nd5* 基因的核苷酸替换速率计算家蚕在 7.1 Ma 与野桑蚕的分化. 这两个研究得到的分化时间远长于化石证据. 因此, 蚕的分化时间仍需要进一步的研究.

人类活动, 如人类的迁徙会影响驯化动物的遗传结构. 人类根据需求培育出了多种表型的驯化动

物. 在培育过程中, 驯化动物的群体大小经常经历扩张. 目前, 家蚕有 1000 多个品系, 它们的表型是多种多样的. 然而, 已有的研究结果则表明, 在驯化过程中, 家蚕的群体大小保持不变^[8,9]. 因此, 需要更多的研究来揭示家蚕在驯化过程中的群体结构变化.

本研究利用线粒体 DNA 和核基因 DNA 序列研究家蚕及其近缘物种的系统发生关系和进化历史. 首先, 相对于日本野桑蚕, 本研究的数据支持以下观点: 中国野桑蚕和家蚕具有更近的亲缘关系. 在放松的分子钟模型下, 利用贝叶斯法和极大似然法计算了三种蚕的分化时间. 研究得到的结果与化石证据和历史记载一致. 结果表明, 家蚕在驯化过程中经历了群体扩张. 这些结果能帮助人们更好地理解蚕的进化历史.

1 材料与amp;方法

1.1 材料收集与 DNA 提取

16 个供试家蚕品种由西南大学家蚕基因库提供, 包括华系、日系、欧系和热带品系等四大地理品系. 13 个中国野桑蚕和 5 个日本野桑蚕分别来自中国和日本的不同地理区域. 系统发生分析中, 用黑点白蚕蛾 (*Ernolatia moorei*)和蓖麻蚕(*Samia cynthia ricini*)作为外类群, 前者与家蚕同属于家蚕蛾科(Bombycidae), 后者属于天蚕蛾科(Saturniidae)(表 1).

取单一供试家蚕和野桑蚕的蛹或幼虫, 用酚-氯仿法提取蚕基因组^[13].

1.2 基因的选择、克隆及测序

本研究选择了 2 个核基因 (Period, *Period*; tyrosine hydroxylase gene, *TH*). 前人研究表明, 鳞翅目昆虫 Period 蛋白 PAS 结构域的氨基酸替换速率比其他核基因编码蛋白质快(是多巴胺脱羧酶 dopa decarboxylase 的 4.9 倍; 是延伸因子 1 α elongation factor-1 α 的 44 倍)^[14]. 因此, 本文认为, 该基因具有较快进化速率, 可用来区分近缘物种. *TH* 基因在家蚕驯化过程中经历了人工选择, 因此该基因在本研究中被用于系统发生分析^[15,16]. 另外, 2 个经典的线粒体基因 *COI*(cytochrome oxidase I)和 *CR*(control region)也被用于本研究的分析.

基于家蚕核基因组和线粒体基因组序列分别设

表 1 家蚕和野桑蚕材料

物种	编号	品系	采集地	化性 ^{a)}
家蚕 <i>B. mori</i>	<i>B. mori</i> S1	诸桂	中国	CV1
	<i>B. mori</i> S2	黄波	中国	CV1
	<i>B. mori</i> S3	四川三眠	中国	CV1
	<i>B. mori</i> S4	黔三眠	中国	CV1
	<i>B. mori</i> S5	印度三眠	印度	T
	<i>B. mori</i> S6	赤熟	日系	JV1
	<i>B. mori</i> S7	欧 18	欧洲	EV1
	<i>B. mori</i> S8	意 16	欧洲	EV1
	<i>B. mori</i> S9	J115	日系	JV2
	<i>B. mori</i> S10	872	日系	JV2
	<i>B. mori</i> S11	大造 N	中国	CVd
	<i>B. mori</i> S13	乌 G	乌克兰	EV2
	<i>B. mori</i> S14	乌 F	乌克兰	EV2
	<i>B. mori</i> S15	柬埔寨	柬埔寨	T
	<i>B. mori</i> S16	白夏 B	中国	CV3
	<i>B. mori</i> S17	C108N	中国	CV2
中国野桑蚕 Chinese <i>B. mandarina</i>	<i>B. man</i> -CHN Y1	湖州	浙江	野生型
	<i>B. man</i> -CHN Y2	南充	四川	野生型
	<i>B. man</i> -CHN Y3	安康	陕西	野生型
	<i>B. man</i> -CHN Y4	宜昌	湖北	野生型
	<i>B. man</i> -CHN Y5	泸州	四川	野生型
	<i>B. man</i> -CHN Y6	澧县	湖南	野生型
	<i>B. man</i> -CHN Y7	苏州	江苏	野生型
	<i>B. man</i> -CHN Y8	青木关	重庆	野生型
	<i>B. man</i> -CHN Y9	荣昌	重庆	野生型
	<i>B. man</i> -CHN Y11	资阳	四川	野生型
	<i>B. man</i> -CHN Y12	彭山	四川	野生型
	<i>B. man</i> -CHN Y13	洪雅	四川	野生型
	<i>B. man</i> -CHN Y14	祥云	云南	野生型
日本野桑蚕 Japanese <i>B. mandarina</i>	<i>B. man</i> -JPN J2	东区	日本福岡	野生型
	<i>B. man</i> -JPN J7	福岡	日本福岡	野生型
	<i>B. man</i> -JPN J227	福岛	日本福岛	野生型
	<i>B. man</i> -JPN J238	静岡	日本静岡	野生型
	<i>B. man</i> -JPN J341	北海道	日本北海道	野生型
黑点白蚕蛾 <i>E. moorei</i>	<i>E. moorei</i>	东南亚	东南亚	外类群
蓖麻蚕 <i>S. c. ricini</i>	<i>S. c. ricini</i>	广西	广西	外类群

a) CV1: 华系一化; CV2: 华系二化; CV3: 华系三化; CVd: 华系多化; JV1: 日系一化; JV2: 日系二化; EV1: 欧系一化; EV2: 欧系二化; T: 热带品系; Wild: 野桑蚕

计这 4 个位点的引物(表 2). *S. c. ricini* 的 *TH* 基因引物是基于 *TH* 的 cDNA 序列设计的^[17]. 对于 *E. moorei* 的 *TH* 基因, 本实验首先下载鳞翅目昆虫 *TH* 基因编码序列, 通过寻找保守区域设计引物, 经 PCR 获得目的产物. PCR 产物经胶回收后克隆到 pMD-19 载体 (TaKaRa, 大连). 随机挑选 2~3 个阳性克隆进行测序, 以避免 PCR 过程产生的假多态性位点^[18].

表 2 基因引物、长度及最适核苷酸替换模型

DNA 类型	位点	引物(5'→3')	长度(bp) ^{a)}	模型
线粒体 DNA	<i>COI</i>	F: GCCCTAAACAACCAAAAG R: GGCGACAAATCATAAAAGATA	718	GTR+I
	<i>CR</i>	F: GCAACTGCTGGCACAAAAT R: TGAGGTATGAGCCCAAAAGC	972	TrN+G
串联线粒体序列			1690	GTR+I
核 DNA	<i>Period</i>	F: ATAAACGAAGACCATTTCGGT R: AATCCTTTTCGTGATGCGTCA BmF ^{b)} AAGTGGTAAGGGTTGAGGGT BmR ^{b)} CGAAGGTTTGTCTTGCTGT	1109	HKY+G
	<i>TH</i>	ErmF ^{c)} TGCATTGGTCTGTTTGCAAT ErmR ^{c)} GAACGACGGCGAGATGGGT SamF ^{d)} CATCAGATTCCGTCATTA SamR ^{d)} GCACCCAATAAGTTCTA	969	TVM+I

a) 比对后序列长度; b) 扩增家蚕及野桑蚕 TH 基因的引物; c) 扩增 *E. moorei* TH 基因的引物; d) 扩增 *S. c. ricini* TH 基因的引物

1.3 序列比对

测序得到的序列用 MUSCLE3.6 进行比对^[19]. 2 个线粒体基因序列及 2 个核基因序列分别被连接成 1 个串联序列.

1.4 系统发生分析

分别用 3 种不同方法研究 3 种蚕的系统发生关系: 极大似然法(maximum likelihood, ML)、最大简约法(maximum parsimony, MP)、贝叶斯法(Bayesian inference, BI). *E. moorei* 和 *S. c. ricini* 的 2 个核基因及 2 个线粒体基因的同源序列被当作外群. 用 PAUP 中的 ILD 法(incongruence length-difference)检测序列碱基组成的异质性, 以探测能否将核基因与线粒体基因合并起来研究^[20,21]. 在 BI 和 ML 分析中, Modeltest Server 1.0 用来检测序列最佳的核苷酸替换模型^[22].

用 PAUP*4.0b10 软件构建最大简约树^[21]. 通过 TBR(tree bisection-reconnection)交换算法的启发式搜索(heuristic search)得到最大简约树, 进行 1000 次 Bootstrap 自展法检验, 空位视为缺失数据. 用 PHYML 3.0 软件的极大似然法构建系统发生树, 根据 Modeltest 检测结果设置参数, 用 BioNJ 树作为起始树, 进行 1000 次 Bootstrap 自展法检验^[23].

用 MrBayes, V3.1.2 软件构建贝叶斯树^[24]. 4 条马尔科夫链(3 条热链和 1 条冷链)独立运行 2 次, 每次运行 1×10⁷ 代, 每 1000 代对系统树抽样 1 次. 用 AWTY 在线程序评估 2 次运行得到的后验概率的收敛性, 以确定 burn-in 值^[25].

用 Network 4.516 软件进行单倍型网络分析, 该分析基于 Median-Joining 模型^[26].

1.5 重组位点检测

用 RDP3 软件检测个体间是否存在 *Period* 基因存在重组^[27]. 本研究用提到的应用软件中的非参数 RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA 及 SISTER SCAN 等不同方法检测重组. 参数采用默认设置.

1.6 分化时间的计算

采用 BEAST v1.5.4 和 r8s v1.7 两种软件计算 3 种蚕的分化时间^[28,29]. 采用不相关松散对数正态分子钟(uncorrelated relaxed lognormal clock)模型, 用串联的 2 个线粒体基因来计算分化时间. BEAST 可以模拟不同群体的历史变化模型, 并在某一特定模型下进行分化时间的计算. 本文使用 3 种不同模型: 群体大小一致模型(constant population size)、群体指数增长模型(exponential population growth)、群体扩张模型(expansion population growth). 在每一种群体变化模型下, 独立运行 2 次, 每次运行 1×10⁷ 代. 用 Tracer v.1.5 分析 BEAST 抽样产生的结果, 取平均值及 95% 最高后验密度区间(95% highest posterior density interval, HPD)内的值作为几种蚕的分化时间. BEAST 需要 1 个或多个已知的校准点作为先验值, 可以用化石证据或已知的数据作为校准点. 在西殷村古代遗址发现半个蚕蛹壳, 该遗址经检测距今约 5000 年^[30]; 在浙江良渚新石器时代文化遗址发现丝线及丝绸残片, 经鉴定该遗址距今也有 5000 年. 这些结果说明, 家蚕的驯化及饲养至少发生在距今 5000 年以前^[31]. 由于中国的古代农业从新石器时代开始^[30], 本文利用 0.004~0.011 Ma 作为校准时间来检测蚕的分化时

间. 另外, Pan 等人^[11]的研究表明, 日本野桑蚕和中国野桑蚕的分化时间大约在 1.11~1.45 Ma, 因此本实验也用此时间作为校准点来计算家蚕与野桑蚕的分化时间.

R8s 软件基于极大似然法计算分化时间. 本实验采用该软件中的非参数速率平滑法(nonparametric rate smoothing NPRS)及半参数罚分似然法(semi-parametric penalized likelihood PL)^[32,33]. 这两种方法均以线粒体基因构建的 ML 树作为起始树. 与 BEAST 相似, 分别利用两个校准点(家蚕与中国野桑蚕的分化时间: 0.004~0.011 Ma; 野桑蚕的分化时间: 1.11~1.45 Ma). 对 NPRS 和 PL 法, 分别选择 Powell 和 Truncated Newton 法进行枝长优化. 分化时间的 95% 置信区间通过 r8s-bootstrap 程序包计算得到^[34]. 这个程序包也需要 Phylip v3.6 内的 SEQBOOT 程序进行随机抽样处理(<http://www.bioinformatics.uthscsa.edu/www/phylip/>).

1.7 群体动力学研究

基于串联线粒体序列, 实验利用 Arlequin 3 中的 2 种方法推测家蚕及中国野桑蚕的群体动力学演化^[35]. 第一种是个体间核苷酸不配对分析(pairwise mismatch distribution); 第二种是用 Fu's F_s 检验来检测稀有碱基是否过量, 稀有碱基过量反应群体可能经历了扩张^[36]. 由于群体较小($N=5$), 没有对日本野桑蚕进行这项分析. BEAST 中 Bayesian skyline plots 法(BSP)同样用来计算蚕的群体动力学历史^[28].

2 结果与分析

2.1 序列特征

本实验采用了 16 个家蚕, 13 个中国野桑蚕和 5 个日本野桑蚕样本, 基本在所有这些样本内克隆并测序了 2 个线粒体基因(其中家蚕样本 *B. mori* S8 和中国野桑蚕样本 *B. man-CHN* Y12 的 CR 片段没有扩增出来)和 2 个核基因的序列(其中家蚕样本 *B. mori* S4, 日本野桑蚕样本 *B. man-JPN* J341 以及外群 *S. c. ricini* 的 *Period* 基因片段没有扩增出来). 对这些序列进行比对, 发现相对于日本野桑蚕, 家蚕和中国野桑蚕具有更高的序列相似性, 并且大部分核苷酸分化发生在日本野桑蚕与中国野桑蚕/家蚕之间. 另外, 日本野桑蚕在线粒体基因及核基因上都含有大

量特异的插入/缺失(indel), 如在线粒体控制区(CR)252 bp 的插入, 在 *Period* 基因内的 2 个插入(285 和 45 bp).

本文所有序列均已递交至 GenBank(GenBank 登录号: *COI*: GQ423211-GQ423244, GU360733, HQ132262; *CR*: GQ423245-GQ423276, GU360735, HQ132261; *Period*: HQ132228-HQ132259, HQ132227; *TH*: GQ423342-GQ423346, GQ415508-GQ415537, GU360734, HQ132260).

2.2 系统发生分析

多位点序列联合分析有利于获得更准确的系统发生关系. 因此, 实验用 ILD 法即同质性检验来检测 2 个线粒体基因、2 个核基因及分析 4 个基因能否串联在一起^[20]. 结果表明, 2 个线粒体基因(*COI*, *CR*)的进化速率相似, 可以合并进行联合分析($P=0.349$). 而 2 个核基因(*Period-TH*, $P=0.001$)及核基因与线粒体基因(nDNA-mtDNA, $P=0.001$)均不能合并. 因此, 本文分别构建了串联线粒体序列、*Period* 基因和 *TH* 基因的系统发生树.

基于线粒体序列, *TH* 基因构建的系统发生树有相似的拓扑结构, 与 *Period* 基因序列构建的系统发生树稍有差异. 因此, 正文仅分析线粒体序列(图 1)和 *Period* 基因构建的系统发生树(图 2). 由图 1 和 2 可见, 系统发生树均包含 2 个主要的支: 一支只包含日本野桑蚕(BJ-clade), 另一支包含全部家蚕品系(在线粒体系统发生树中, 由于 *B. mori* S8 和 *B. man-CHN* Y12 样本 *CR* 序列缺失, 所以这两个样本未用于分析)及中国野桑蚕(B.BC-clade), 表明家蚕和中国野桑蚕具有较近的亲缘关系. 但是, 在用不同基因序列构建系统发生树时, B.BC-clade 内部的拓扑结构有些不同. 在 *Period* 树中, 所有家蚕品系和 3 个中国野桑蚕样本(*B. man-CHN* Y3, *B. man-CHN* Y4 和 *B. man-CHN* Y12)聚类在一起(图 2). 而在线粒体及 *TH* 树中, 所有家蚕先聚类在一起形成一枝(B clade), 然后再和中国野桑蚕聚类在一起(图 1 和网络版附图 1). 另外, 本研究用 Li 等人^[8]测得的其他家蚕品系全线粒体序列中的 *CR* 和 *COI* 位点与我们测得的相应序列共同构建系统发生树, 得到与图 1 相同的结果(网络版附图 2). 虽然用不同基因构建的系统发生树拓扑结构有所不同, 但所有分析均表明, 相对于日本野桑蚕, 家蚕与中国野桑蚕具有更近的亲缘关系, 家蚕由

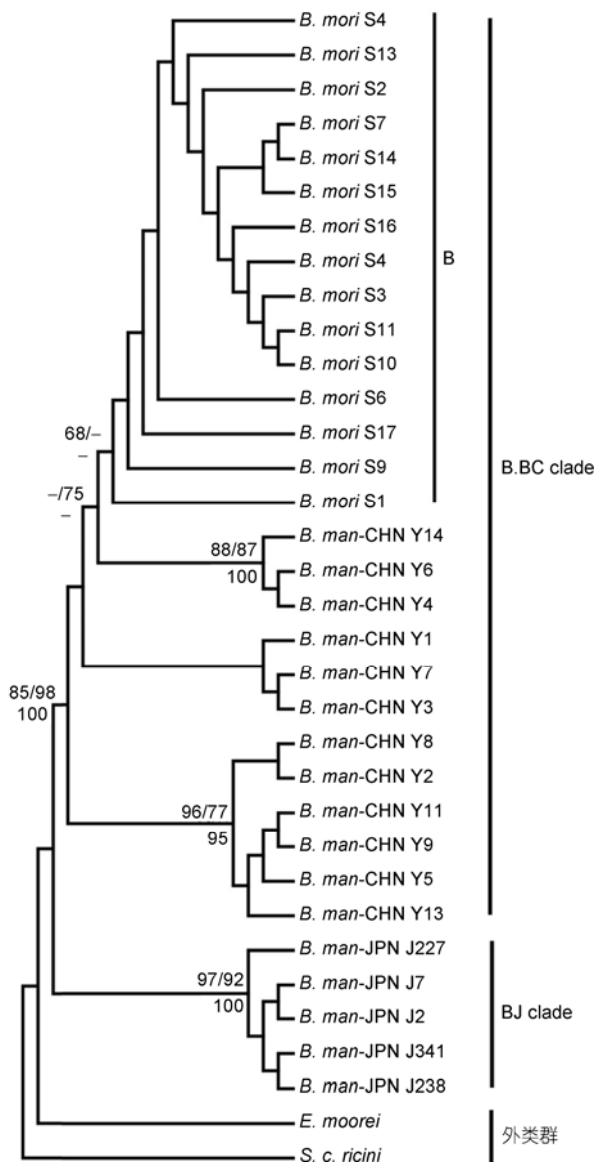


图 1 基于串联线粒体序列的 MP 系统发生树

进化支左上侧值: ML 树的 bootstrap 值; 右上侧值: MP 树的 bootstrap 值; 下侧值: BI 树的后验概率; *B. man-CHN*: 中国野桑蚕; *B. man-JPN*: 日本野桑蚕

中国野桑蚕驯化而来。

2.3 分化时间的计算

用 2 种不同方法: 贝叶斯法(BEAST)和极大似然法(r8s)计算 3 种蚕的分化时间。根据历史和化石证据, 用家蚕与中国野桑蚕的分化时间(0.004~0.011 Mya)作为计算蚕分化时间的第一个校准点(表 3, 灰色背景)。在 BEAST 软件中, 分别用 3 种不同的群体

动力学参数模型计算蚕的分化时间。由表 3 可见, 结果都较为相似。因此, 正文只展示了在群体大小一致的模型下计算的分化时间图谱(图 3)。家蚕和中国野桑蚕最近共同祖先(most recent common ancestor MRCA)出现在 4100 年前。中国野桑蚕与日本野桑蚕约在 23600 年前(Years ago, Ya)分开, 95%最高后验密度区间(95% HPD)的上限和下限分别为 5200 和 50200 年前, 这个时间远早于蚕的驯化时间。群体指数增长模型(1400 Ya(95% HPD: 4000~28300 Ya))及群体扩张模型(20900 Ya(95% HPD: 4100~39100 Ya), 也证实了野桑蚕的分化时间(表 3)。另外, 不同品系的家蚕的扩张时间约在 2000 Ya(95% HPD: 600~4300 Ya), 与另外 2 个模型的结果相似。r8s 内两种极大似然法(PL, NPRS)也分别用于分化时间的分析(表 3, 灰色背景)^[32,33]。整体来看, r8s 得到的结果与 BEAST 的结果相似。NPRS 分析中, 家蚕的传播时间约在 1540 Ya(95%置信区间: 716~2811 Ya), 中国野桑蚕与日本野桑蚕的分化时间为 46373 Ya(95%置信区间: 15931~65743 Ya); PL 分析中前者为 2891 Ya(95%置信区间: 2085~3605 Ya), 而后者为 53798 Ya(95%置信区间: 21642~78438 Ya)。总体上看, 用 2 种软件计算的家蚕传播时间及野桑蚕分化时间都相似, 并符合历史及化石证据^[30,31,37]。

Pan 等人^[11]计算了野桑蚕的分化时间, 约为 1.11~1.45 Ma。因此, 本实验用该数据作为第二个校准点计算蚕的分化时间(表 3, 白色背景)。BEAST 软件中, 同样用 3 种不同群体动力学参数模型计算蚕的分化时间。如表 3 所示, 家蚕的驯化时间约在 286800 Ya(95% HPD: 137900~473900), 家蚕的传播时间为 117100 Ya(95% HPD: 46800~217200 Ya)。R8s 也用第二个校准点计算蚕的分化时间(表 3, 灰色背景)。NPRS 分析中, 家蚕的驯化时间为 750000 Ya(95%置信区间: 586919~818598), 传播时间约在 471667 Ya(95%置信区间: 400933~542400); PL 分析中前者为 410000 Ya(95%置信区间: 331947~544417), 而后者为 400000 Ya(95%置信区间: 236644~659218)。以上所有记过均表明, 用第二个校准点计算出的家蚕驯化时间明显早于化石证据及历史记载^[4,30,31]。因此, 本文认为, Pan 等人^[11]得到的中国野桑蚕与日本野桑蚕的分化时间不合理。

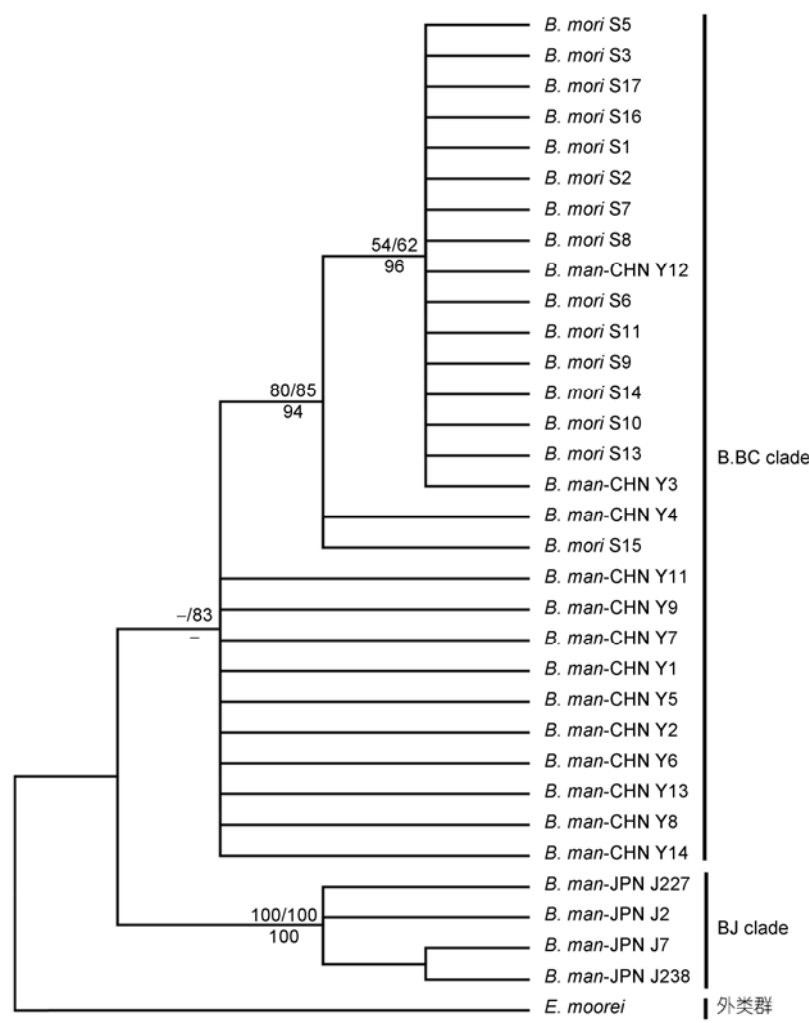


图 2 基于 *Period* 基因序列的 MP 系统发生树

进化枝左上侧值: ML 树的 bootstrap 值; 右上侧值: MP 树的 bootstrap 值; 下侧值: BI 树的后验概率; *B. man-CHN*: 中国野桑蚕; *B. man-JPN*: 日本野桑蚕

2.4 群体动力学分析

基于单倍型数据构建的 median-joining 法网络结构能提供直接的物种进化历史图谱. 图 4A 大图为用串联线粒体 *CR-COI* 序列(包括本研究所用样本所获序列及 Li 等人^[8]所获其他品系蚕的 *CR-COI* 序列)构建的网络结构图. 结果显示, 家蚕和 2 种野桑蚕分别形成了 3 个单倍型群. 与系统发生分析结果一致, 中国野桑蚕与家蚕的亲缘关系较近. 与野桑蚕群相比, 家蚕的单倍型群呈现“星状”分布, 有 1 个高频分布的单倍型. 前人对其他驯化动物, 如牛和羊的群体动力学分析也发现星状分布的单倍型群, 并指出在这些动物的驯化过程中群体经历了扩张^[38,39]. 因此, 本

研究得到的这种放射性结构暗示, 家蚕在进化过程中可能经历了群体扩张^[26,40]. 另外, 本实验利用 Li 等人^[8]获得的蚕及中国野桑蚕的全线粒体基因组序列构建单倍型网络结构(图 4A, 小图), 结果显示, 家蚕的单倍型之间亲缘关系很近, 由于单倍型频率高 (Haplotype diversity, Hd=1.0), 没有像 *CR-COI* 得到的高频率分布的单倍型. 但是, 结果仍表明, 家蚕的单倍型群有呈现“星状”分布的趋势. 以上结果表明, 在进化过程中家蚕群体可能经历了扩张, 而前人研究表明, 家蚕的群体大小在进化过程中保持不变^[8,9]. 因此, 本文将用更为精确的方法计算蚕的群体动力学演化. 另外, 基于 *Period* 基因构建的系统发生树显

表 3 两种算法计算出的分支节点的分化时间^{a)}

软件位点	模型/方法	分化时间(Ya)(95% HPD)		
		<i>B. man</i> -JPN vs. <i>B. man</i> -CHN	<i>B. man</i> -CHN vs. <i>B. mori</i>	<i>B. mori</i> vs. <i>B. mori</i>
R8s mtDNA	NPRS	46373(15931~65743)	5160 [*] (3686~8114)	1540(716~2811)
	PL	53798(21642~78438)	4530 [*] (2142~6810)	2891(2085~3605)
	NPRS	6100000 [*] (4701297~9066110)	750000(586919~818598)	471667(400933~542400)
	PL	5000000 [*] (4413357~6273786)	410000(331947~544417)	400000(236644~659218)
Beast mtDNA	Constant size	23600(5200~50200)	4100 [*] (-)	2000(600~4300)
	Exponential growth	14700(4000~28300)	4400 [*] (-)	2100(600~4300)
	Expansion growth	20900(4100~39100)	4000 [*] (-)	2300(500~4000)
	Constant size	1258800 [*] (1111800~1428500)	286800(137900~473900)	117100(46800~217200)
	Exponential growth	1258900 [*] (1110200~1423500)	352500(138600~620700)	159600(42000~304200)
	Expansion growth	1276700 [*] (1111900~1431200)	311200(130400~492600)	144600(45000~218800)
Beast TH	Constant size	13000(4000~24700)	1300 [*] (300~2600)	900(200~1700)
	Exponential growth	9700(4000~18600)	1600 [*] (400~3500)	1100(200~2510)
	Expansion growth	10700(4000~25900)	1000 [*] (400~3100)	700(200~2100)

a) *: 校准点; 灰色背景内的值是通过第一个校准点(0.004~0.011 Ma)计算出来的结果; 白色背景内的值是通过第二个校准点(1.11~1.45 Ma)计算出来的结果

示, 3 个中国野桑蚕样本与家蚕混合在一起. 本文用单倍型网络结构分析它们是否共享单倍型. 如图 4B 所示, 日本野桑蚕具有相似的单倍型, 并与家蚕及中国野桑蚕的差异较大. 系统发生分析中混合的 3 个中国野桑蚕与家蚕共享单倍型.

实验用 Arlequin v3.0 软件中核苷酸不配对分析及 Fu's F_s 检验来检测家蚕和中国野桑蚕的群体结构^[35]. 核苷酸不配对分析及 Fu's F_s 检验是检验群体结构动力学的经典方法, 前者指不同单倍型之间碱基不配对的相对频率分布, 当种群群体经历扩张时核苷酸不配对分布呈现单峰或平滑分布^[41]; 后者对群体内突变过量较敏感, 可用来检测近期群体的扩张事件, 当 F_s 值显著负时暗示存在群体扩张事件^[42]. 由于群体数目太小($N=5$), 对日本野桑蚕没有进行这项研究. 如图 5A 所示, 家蚕的不配对分析结果显示, 种群单倍型碱基差异分布呈单峰分布, Fu's F_s 值($F_s=-3.769$, $P=0.003$)显著负, 表明家蚕在驯化过程中经历了群体扩张^[41,42]. 另外, Tajima's D 值($D=-1.945$, $P<0.05$)及 Fu 和 Li's D 值($D=-2.535$, $P<0.05$)均显著负也证实了家蚕群体扩张的存在^[36,40]. 中国野桑蚕的核苷酸不配对分析中无明显峰出现, 而 Fu's F_s 值($F_s=-2.252$, $P=0.110$), Tajima's D 值($D=-0.560$, $P>0.1$)及 Fu 和 Li's D 值($D=-0.776$, $P>0.1$)均不显著负, 表明中国野桑蚕群体大小在进化过程中无明显变化. BEAST 软件中的 Bayesian skyline plots 法(BSP)同样用来推算家蚕群体随时间的变化历程, 结果显示, 距今约

1000 Ya 时, 家蚕的群体产生了 1 次扩张(图 5B 虚线处)^[43]. Li 等人^[8]用第二代测序法测得了 29 个家蚕品系及 11 个中国野桑蚕的全线粒体基因组序列, 本文也用这些数据分析家蚕及中国野桑蚕的群体变化历程. 如图 5C 所示, 与串联线粒体序列的核苷酸不配对分析结果相似, 家蚕在第 15 个核苷酸不配对差异处有 1 个平滑的单一的峰, 且 Fu's F_s 值显著负($F_s=-19.131$, $P<0.001$), 表明家蚕经历了群体扩张. BSP 分析同样表明, 家蚕在距今约 1000 年时群体开始扩增(图 5D). 基于 11 个中国野桑蚕全线粒体基因组序列的核苷酸不配对分析仍没有单峰出现, 且 Fu's F_s 值不显著负($F_s=-0.007$, $P=0.321$), 说明中国野桑蚕没有经历群体扩张(图 5C).

3 讨论

3.1 家蚕起源于中国野桑蚕

家蚕和中国野桑蚕在测序获得的 2 个线粒体基因和 2 个核基因上均有较高的序列相似性, 而日本野桑蚕有很多特异的序列特征, 如日本野桑蚕在线粒体 CR 区有 1 个长 216 bp 片段的插入, 在 *Period* 基因内有 2 个独特的插入, 与之前只在日本野桑蚕的 *Arylphorin* 基因中发现 1 个转座子的插入^[5]的研究结果相似. 这些结果表明, 本实验选择的分子标记有足够的核苷酸多态性来区分近缘物种的系统发生关系. 并且从序列特征看, 家蚕与中国野桑蚕的亲缘关系更近.

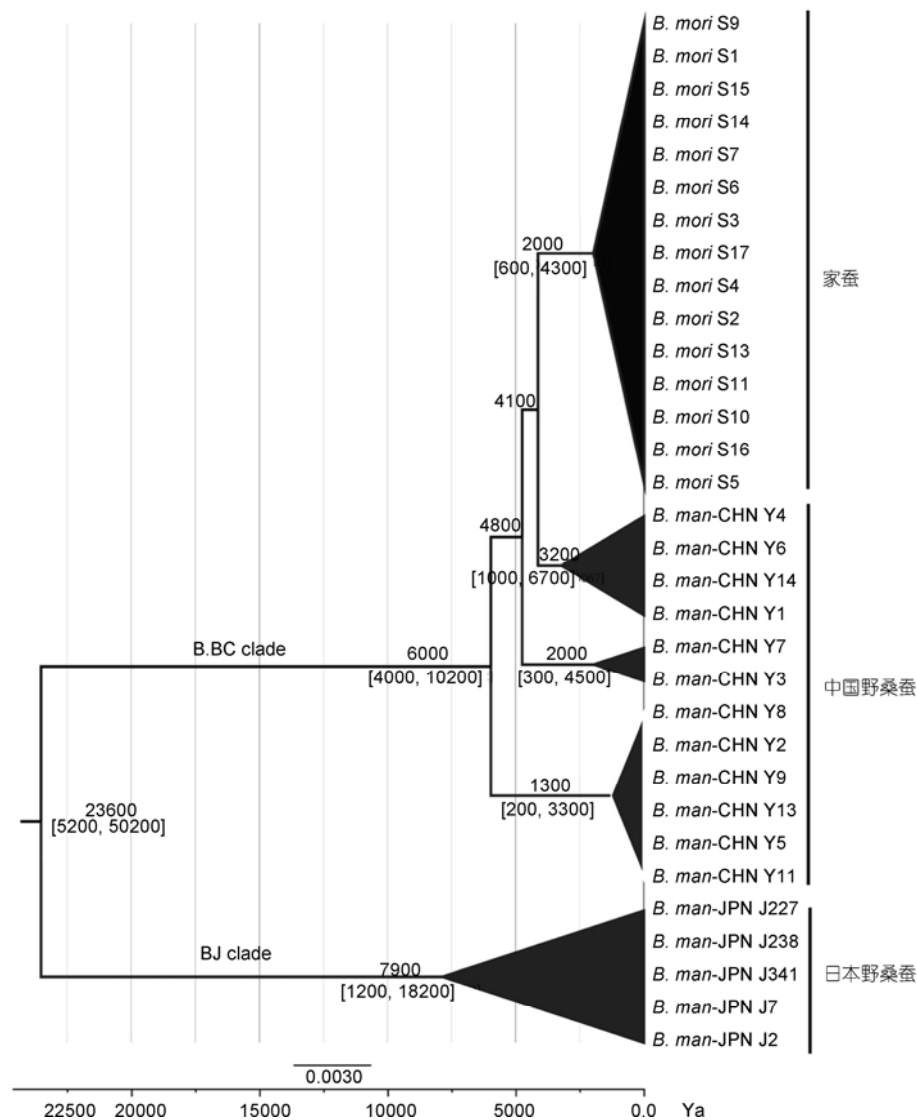


图3 BEAST 软件计算的蚕的分化时间图谱

分支上值: 节点处平均分化时间; 分支下侧括号内值: 95%最高后验密度区间值

基于串联线粒体序列及 2 个核基因序列分别构建的系统发生树得到相似的拓扑机构. 结果表明, 日本野桑蚕首先和另外 2 种蚕分开(图 1), 而后家蚕与中国野桑蚕聚类在一枝上(图 1, B.BC clade). 该结果与之前用核基因组及线粒体基因组全序列得到的结果一致^[8,9]. 用不同位点构建系统发生树时, B.BC clade 内枝的拓扑结构会有所不同. 对于串联线粒体序列和 *TH* 序列, 所有家蚕先聚类在一起(B clade), 而后再和中国野桑蚕聚为一枝. 但是, *Period* 序列构建的系统发生树则显示, 这两类蚕为混合聚类, 不能

将家蚕与中国野桑蚕区分开; 单倍型网络分析也表明, 部分中国野桑蚕与家蚕有一致的单倍型(图 4). 之前有研究发现, 重组会对系统发生分析的准确性产生很大影响, 本实验用 RDP3 软件内多种重组检测方法检验 *Period* 基因是否经历个体间的重组^[44-46]. 结果发现, 家蚕与中国野桑蚕在 *Period* 基因上无重组发生. Lau 等人^[47]在研究家马和野马的系统发生关系时, 也得到了与 *Period* 基因构建的系统发生树相似的结果——家马与野马混合在一起, 并指出这可能由于种质渗入或 2 个物种在研究的基因共享祖先

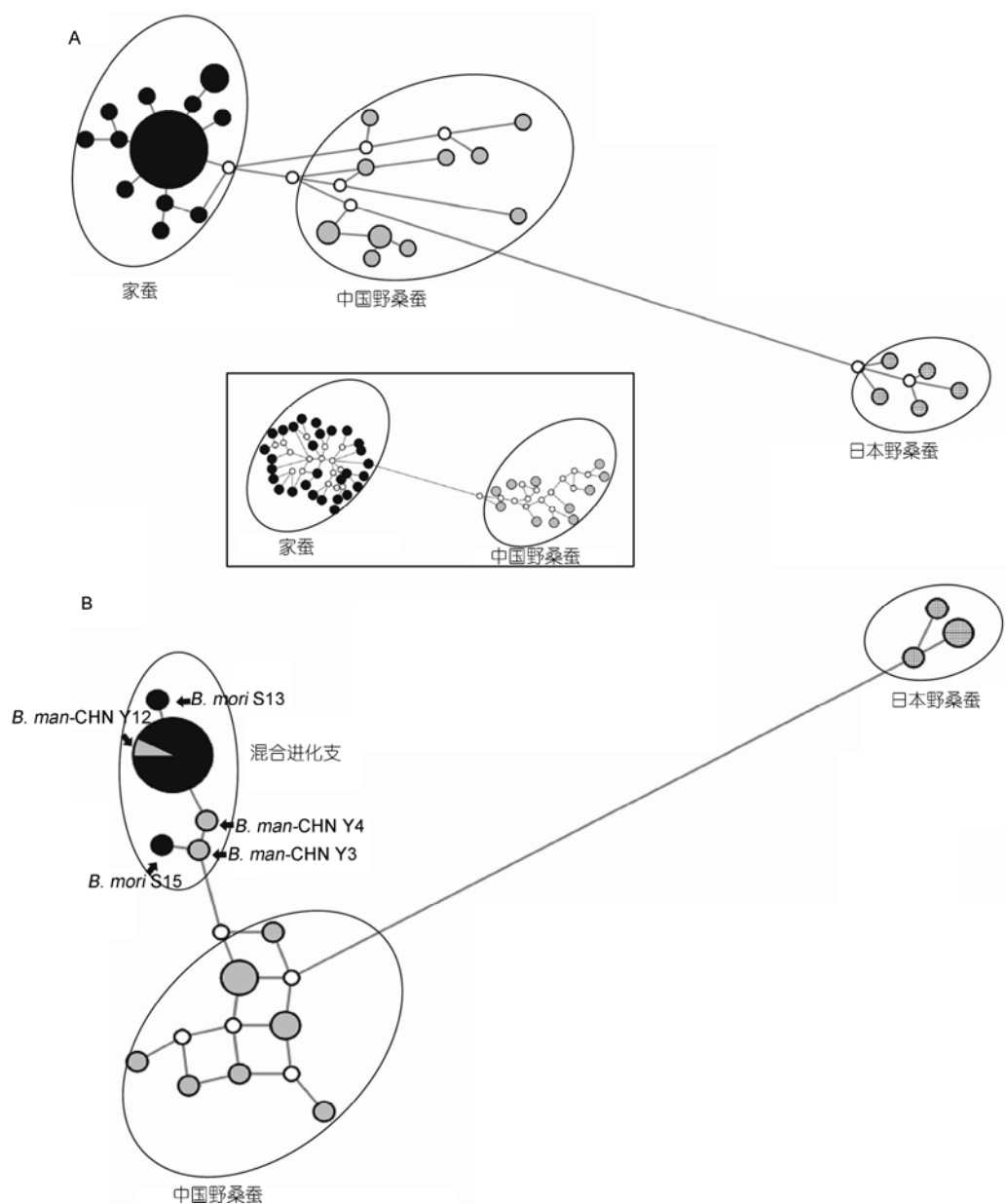


图 4 线粒体序列和 *Period* 基因序列单倍型网络图

A: 粒体序列单倍型网络图(大图为基于串联 *CR-COI* 序列构建的单倍型网络图; 小图为基于全线粒体基因组序列构建的单倍型网络图);
B: *Period* 基因序列单倍型网络图

多态性. Klein 等人^[48]指出, 当分析的 2 个物种的分化时间少于 1~2 百万年时, 祖先多态性会对系统发生树的准确性产生很大影响. 本研究计算的家蚕与野桑蚕分化的时间仅为 4100 年, 所以本文推测家蚕和中国野桑蚕在 *Period* 位点共享祖先多态性是导致这 2 个物种在系统发生分析中不能分开的原因.

3.2 蚕的进化历史

考古学及细胞遗传学研究表明, 家蚕最先在距今 4600 年前由中国野桑蚕驯化而来^[4]. 大约公元前 200 年, 随着人口迁徙, 家蚕及养蚕技术由中国传至韩国和日本^[49]. 汉朝(公元前 206~公元前 220 年)的使者开通了丝绸之路, 增加了古亚洲和古欧洲的联系,

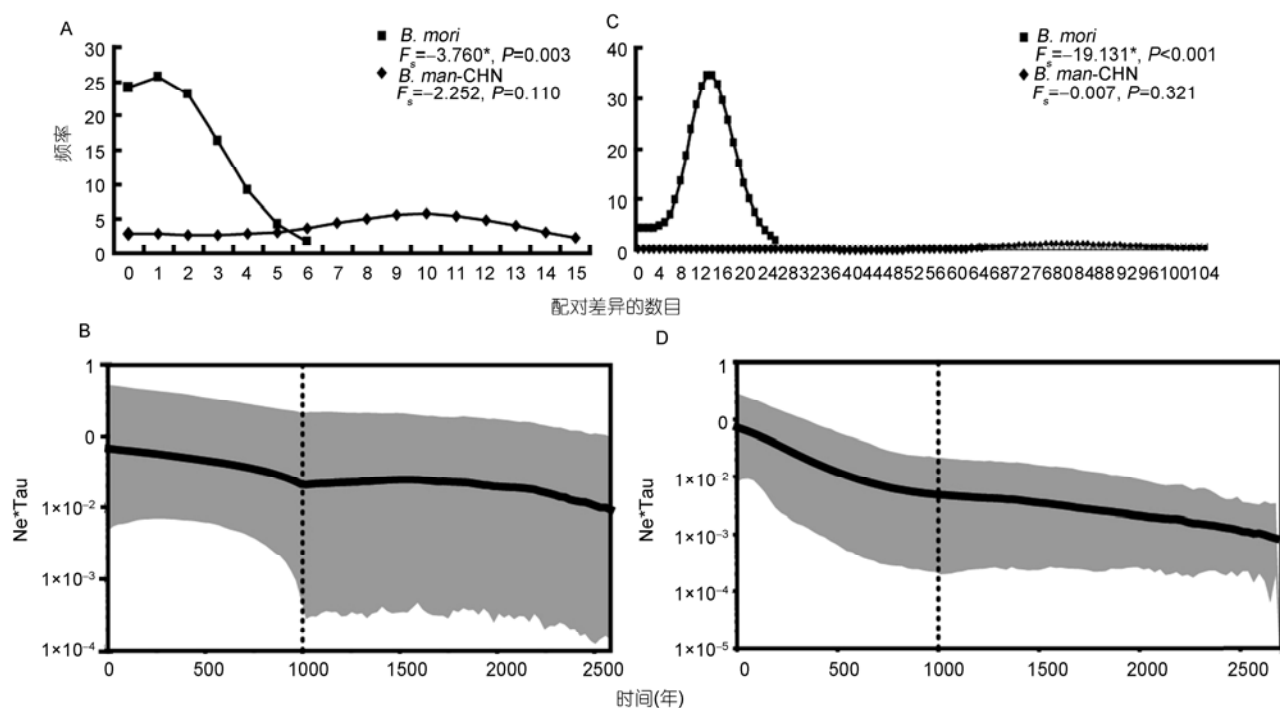


图5 串联线粒体序列和全线粒体基因组序列的核苷酸不配对分布和 Bayesian skyline plots 分析

A 和 B: 串联线粒体序列的核苷酸不配对分析和 Bayesian skyline plots 分析; C 和 D: 全线粒体基因组序列的核苷酸不配对分析和 Bayesian skyline plots 分析; A 和 C 图中正方形曲线表示家蚕, 菱形曲线表示中国野桑蚕

为不同国家之间的贸易及文化交流提供了便利。在这期间, 养蚕技术及丝绸也从中国经中亚国家传至欧洲^[10]。因此, 家蚕由中国传至其他国家的历史在一定程度上代表了古代国家之间贸易和文化交流的历史。尽管很多书籍记载了这段历史, 但分子水平的证据仍然匮乏。如前文所述, Pan 等人^[11]计算出中国野桑蚕与日本野桑蚕的分化时间为 1.11~1.45 Ma, 家蚕的驯化时间为 1.08~1.41 Ma, 远远早于人类农业起源的历史^[30]。Yukuhiro 等人^[12]计算出家蚕和日本野桑蚕约在 7.1 Ma 分开。两个研究得到的结果相差巨大, 可能是他们用的核苷酸替换速率不同所致。Pan 等人^[11]用的是鳞翅目昆虫凤蝶科(Papilionidae)线粒体基因 COI 和 COII 的核苷酸替换速率, 而 Yukuhiro 等人^[12]用的是鞘翅目昆虫步甲科(Carabidae)线粒体 *Nd5* 基因的核苷酸替换速率。这两个科的昆虫与蚕的亲缘关系较远。这两篇文章均用直线回归法(linear regression method)计算分化时间, 该方法假设在进化过程中核苷酸替换速率保持不变^[28]。然而, 这种简单的模型是不现实的, 因为替换速率在种内及种间是不断变化的^[50]。

本研究利用 2 种不同的放松分子钟模型计算蚕的分化时间。一种基于贝叶斯蒙特卡洛-马尔科夫原理, 且在计算过程中可模拟群体变化的模型。另一种基于极大似然原理。表 3 结果表明, 极大似然法得到的结果稍大于贝叶斯蒙特卡洛-马尔科夫原理得到的结果, 但总体上较为相似。家蚕与中国野桑蚕的分化时间约为 4100 年, 与化石证据一致^[4]。总体上看, 家蚕的驯化时间晚于其他物种, 如家狗、水稻等。Gupta^[51]指出, 家狗作为人类狩猎的帮手, 是第一个被驯化的物种, 而水稻作为食物也被较早驯化, 这些对原始人类的生存至关重要。而丝绸并不是原始人类必不可少的, 所以家蚕的驯化要晚一些^[51]。家蚕有 4 个主要地理品系, 它们的分开时间为 2000 Ya(95% HPD: 600~4300 Ya), 本文计算的时间也与历史记载相似, 大约 2000 年前家蚕由中国传至日本和其他国家^[37,49]。以上结果表明, 本文计算的蚕的分化时间是可靠的。中国野桑蚕与日本野桑蚕分开的时间为 23600 Ya(95% HPD: 5200~50200 Ya), 稍早于日本群岛完全从亚洲大陆分开的时间(约 20000 年前)^[51]。如前所述, 中国野桑蚕与日本野桑蚕的染色体数目不

同, 并有研究表明, 朝鲜半岛北部地区可能是两种野桑蚕的分界区域^[52,53]. 因此, 加上来源于朝鲜半岛的野桑蚕样本有利于野桑蚕分化时间的计算. 遗憾的是, 目前很难得到该区域的野桑蚕样本. 但是, 前人的很多研究发现, 来自朝鲜半岛的野桑蚕与日本野桑蚕拥有相似特征, 如相同的染色体数目、*Arylphorin* 基因都含有转座子的插入. 所以, 本文认为, 添加朝鲜半岛的野桑蚕样本可能对计算分化时间的准确性影响不大. 本实验同样用核基因 *TH* 计算蚕的分化时间(表 3), 结果发现, 所有时间都小于用线粒体序列计算的时间. 由于 BEAST 软件基于系统发生树框架推演进化时间, 线粒体序列与 *TH* 基因构建的系统发生树的拓扑机构的不同可能对分化时间的计算有影响. Yu 等人^[16]发现, *TH* 基因在驯化过程中经历了很强的人工选择, 导致 DNA 多态性降低, 也有可能影响到分化时间的计算.

基于化石证据, 用 BEAST 软件计算出蚕线粒体基因 *COI* 和 *CR* 的核苷酸替换速率约为 1.024×10^{-6} (95% HPD: $0.358 \times 10^{-6} \sim 1.872 \times 10^{-6}$)/site/year. 此替换速率与某些哺乳动物(如人(*Homo sapiens*): $0.164 \times 10^{-6} \sim 1.398 \times 10^{-6}$ /site/year^[54]; 阿德莱企鵝(*Pygoscelis adelia*): $0.53 \times 10^{-6} \sim 1.43 \times 10^{-6}$ /site/year^[55]; 野牛(*Bos primigenius*): $0.15 \times 10^{-6} \sim 1.30 \times 10^{-6}$ /site/year^[56])相近, 但远高于其他昆虫(如暗果蝇(*Drosophila obscura*): $2.0 \times 10^{-8} \sim 2.9 \times 10^{-8}$ /site/year^[57]; 凤蝶(*Papilio*): $7.8 \times 10^{-9} \sim 10.2 \times 10^{-9}$ /site/year^[58]; 意大利蜜蜂(*Apis mellifera*): $3.4 \times 10^{-8} \sim 4.9 \times 10^{-8}$ /site/year^[59]). Ho 等人^[60]指出, 计算分化时间较短物种的替换速率时会依赖时间尺度(time scale dependent), 许多因素会影响替换速率的计算, 如测序错误、负选择、校准点的选择、突变热点. 对驯化物种来说, 驯化通常在较短时间内完成, 因此校准点的选择可能导致过高估计替换速率^[61,62]. 本研究用化石证据作为校准点, 但由于化石推断的时间只表示距今最近的时间^[63], 所以用中国农业起源时间作为校准点的上限^[30]. Ho 和 Phillips^[64]指出, 采用化石数据并给定一个时间上限作为校准点对分化时间的计算有一定益处. 然而, 本研究利用的校准点仍然对过高的蚕线粒体替换速率有影响. 家蚕线粒体基因经历了非常强的负选择(Tajima's *D*: -1.945, $P < 0.05$)也对核苷酸替换速率的计算有影响. 尽管如此, 本研究计算的蚕的替换速率将对家蚕蛾科(Bombycidae)及相近物种的系统发生分析提供帮助.

目前, 家蚕在全世界范围内有超过 1000 个品系, 这些品系通常有多种多样的表型, 如不同体色、不同化性等. 对其他驯化动物而言, 人类活动会影响它们的群体结构^[39]. 在驯化过程中, 驯化动物的群体经常会经历扩张, 如家牛、山羊及家猪都经历了群体扩张^[38,65]. 然而, Xia 等人^[9]及 Li 等人^[8]指出, 家蚕的群体大小在驯化过程中保持不变, 这种与其他驯化动物不同的群体结构变化很难被理解. 因此, 本文认为, 需要更多研究来重塑家蚕的群体动力学历史. 本文采用了 4 种不同方法: 单倍型网络分析、核苷酸不配对分析、 F_s 检测及 Bayesian skyline plots 检测. 所有结果都显示, 家蚕在距今约 1000 年左右经历了一次群体扩张, 而中国野桑蚕的群体大小则保持不变(图 4, 5A 和 B). 用 Li 等人^[8]测序得到的 29 个家蚕和 11 个中国野桑蚕的全线粒体基因组序列做相似的分析, 发现家蚕群体扩张的信号, 而这与 Li 等人的研究结果相反, 他们认为在进化过程中家蚕的群体保持不变, 但是中国野桑蚕群体大小不变(图 5C 和 D). 以上结果提示, 家蚕与其他驯化动物相似, 在进化过程中经历了群体扩张. 本研究的结果与 Li 等人^[8]和 Xia 等人^[9]的结果不同, 但原因尚不清楚. 一个可能的原因是 Xia 等人^[9]在进行群体动力学模拟时用了—个很简单的模型, 该模型并未考虑瓶颈效应的影响. 不同研究得到相反结果暗示, 家蚕的群体动力学演化历程还需进一步的分析和探究. 如上所述, 驯化动物的群体结构会随着人类的需求有所变化, 如对毛色的喜好、肉质的要求等^[38,39,65]. 对家蚕来说, 人类的需求, 如对丝品质的要求, 也会影响家蚕的遗传结构. 因此, 本文推测, 家蚕群体经历扩张的模型是合理的. 实验检测出, 家蚕群体扩张大约发生在距今 1000 年前的中国宋朝(960 A.D.~1279 A.D.). 在宋朝前, 中国的首都及经济中心大都在中国北部, 随着南宋(1127 A.D.~1279 A.D.)首都迁移至中国南方, 人口也随之迁徙至南方, 并携带有先进的农业技术^[66]. 在这个时期, 北方家蚕品系传播至南方后需要适应新的环境, 如北方比南方冷, 且两个地方的家蚕品系的杂交同样会改变家蚕的遗传结构. 另外, 南宋时期农业技术发生了巨大创新, 此时期被称为“绿色革命”, 先进的农业技术对家蚕品质的改良也有推动作用^[67]. 所有结果表明, 家蚕群体在距今约 1000 年时的扩张可能与宋朝人口的迁徙及农业技术的改进有关.

致谢 感谢杜周和博士在野桑蚕样本的采集, 代方银博士在家蚕样本的采集方面的帮助, 感谢本小组所有成员的帮助, 感谢东京大学 Hirohisa Kishino 教授在数据分析上的建议. 本研究所用外群 *Ernolatia moorei* 由 Stefan Naumann 采集, 并由 Charles Mitter and Jerome Regier 提供, 他们的基金来源于美国国家科学基金(NSF—USA No. 0531769, 0531626, DEB-0212910).

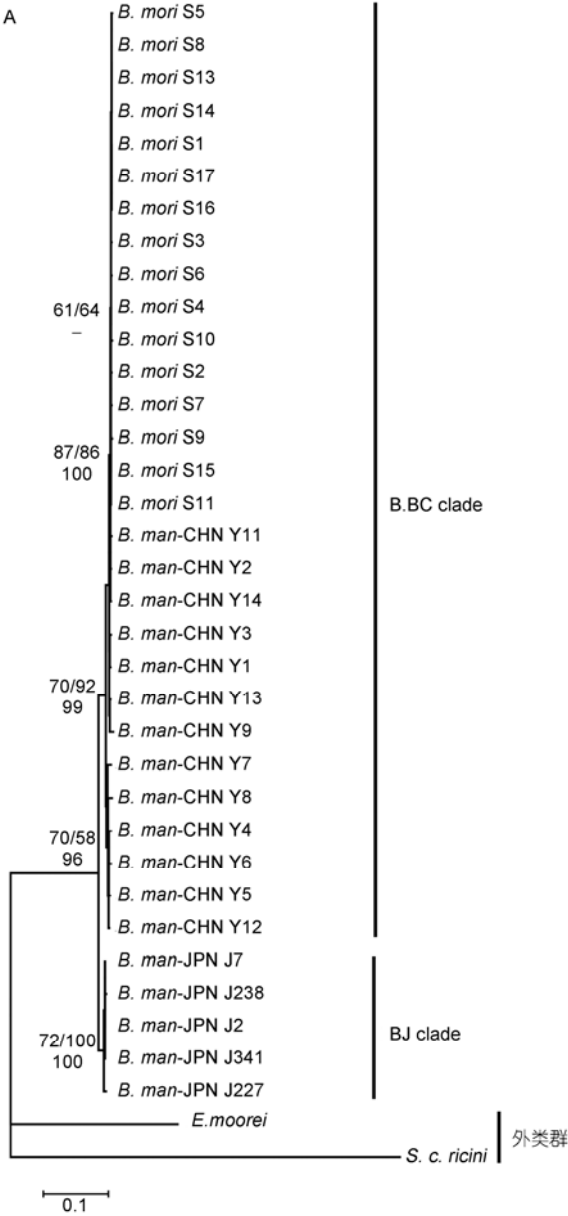
参考文献

- 1 Astaurov B L, Rovinskaya I S. Chromosome complex of Ussuri geographical race of *Bombyx mandarina* M. with special reference to the problem of the origin of the domesticated silkworm, *Bombyx mori*. Cytology, 1959, 1: 327–332
- 2 Banno Y, Nakamura T, Nagashima E, et al. M chromosome of the wild silkworm, *Bombyx mandarina* ($n=27$), corresponds to two chromosomes in the domesticated silkworm, *Bombyx mori* ($n=28$). Genome, 2004, 47: 96–101
- 3 Chikushi H. Genes and genetical stocks of the silkworm. Tokyo: Keigaku Pub. Co., 1972
- 4 Yoshitake N. Phylogenetic aspects on the origin of Japanese race of the silkworm, *Bombyx mori* L. J Sericol Sci Japan, 1968, 37: 83–87
- 5 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
- 6 Arunkumar K P, Metta M, Nagaraju J. Molecular phylogeny of silkmoths reveals the origin of domesticated silkmoth, *Bombyx mori* from Chinese *Bombyx mandarina* and paternal inheritance of *Antheraea proylei* mitochondrial DNA. Mol Phylogenet Evol, 2006, 40: 419–427
- 7 Li A, Zhao Q, Tang S, et al. Molecular phylogeny of the domesticated silkworm, *Bombyx mori*, based on the sequences of mitochondrial cytochrome *b* genes. J Genetics, 2005, 84: 137–142
- 8 Li D, Guo Y, Shao H, et al. Genetic diversity, molecular phylogeny and selection evidence of the silkworm mitochondria implicated by complete resequencing of 41 genomes. BMC Evol Biol, 2010, 10: 81
- 9 Xia Q, Guo Y, Zhang Z, et al. Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). Science, 2009, 326: 433–436
- 10 Liu X. The Silk Road in World History. USA: Oxford University Press, 2010
- 11 Pan M, Yu Q, Xia Y, et al. Characterization of mitochondrial genome of Chinese wild mulberry silkworm, *Bombyx mandarina* (Lepidoptera: Bombycidae). Sci China Ser C-Life Sci, 2008, 51: 693–701
- 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 Nagaraja G M, Nagaraju J. Genome fingerprinting of the silkworm, *Bombyx mori*, using random arbitrary primers. Electrophoresis, 1995, 16: 1633–1638
- 14 Regier J C, Fang Q Q, Mitter C, et al. Evolution and phylogenetic utility of the *period* gene in Lepidoptera. Mol Biol Evol, 1998, 15: 1172–1182
- 15 Futahashi R, Sato J, Meng Y, et al. yellow and ebony are the responsible genes for the larval color mutants of the silkworm *Bombyx mori*. Genetics, 2008, 180: 1995–2005
- 16 Yu H S, Shen Y H, Yuan G X, et al. Evidence of selection at melanin synthesis pathway loci during silkworm domestication. Mol Biol Evol, 2011, 28: 1785–1799
- 17 Hashimoto K, Yamano Y, Morishima I. Induction of tyrosine hydroxylase gene expression by bacteria in the fat body of eri-silkworm, *Samia cynthia ricini*. Comp Biochem Phys B, 2008, 149: 501–506
- 18 Hilton H, Gaut B S. Speciation and domestication in maize and its wild relatives: evidence from the globulin-1 gene. Genetics, 1998, 150: 863–872
- 19 Edgar R C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res, 2004, 32: 1792
- 20 Farris J S, Källersjö M, Kluge A G, et al. Testing significance of incongruence. Cladistics, 1994, 10: 315–319
- 21 Swofford D L. PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4. 2003
- 22 Posada D. ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. Nucleic Acids Res, 2006, 34: 700–703

- 23 Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*, 2003, 52: 696–704
- 24 Ronquist F, Huelsenbeck J P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 2003, 19: 1572–1574
- 25 Nylander J A, Wilgenbusch J C, Warren D L, et al. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, 2008, 24: 581–583
- 26 Bandelt H J, Forster P, Rhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, 1999, 16: 37–48
- 27 Martin D P, Lemey P, Lott M, et al. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, 2010, 26: 2462–2463
- 28 Drummond A J, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*, 2007, 7: 214
- 29 Sanderson M J. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, 2003, 19: 301–302
- 30 Underhill A P. Current issues in Chinese neolithic archaeology. *J World Prehistory*, 1997, 11: 103–160
- 31 顾国达, 徐俊良. 论我国蚕丝业的多中心起源. *浙江大学学报(人文社会科学版)*, 2003, 33: 42–48
- 32 Sanderson M J. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol*, 1997, 14: 1218–1231
- 33 Sanderson M J. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol Biol Evol*, 2002, 19: 101–109
- 34 Eriksson T. r8s Bootstrap Kit. Stockholm: Bergius Botanic Garden, 2002
- 35 Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform*, 2005, 1: 47–50
- 36 Fu Y X, Li W H. Statistical tests of neutrality of mutations. *Genetics*, 1993, 133: 693–709
- 37 Goldsmith M R, Shimada T, Abe H. The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol*, 2005, 50: 71–100
- 38 Bradley D G, MacHugh D E, Cunningham P, et al. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci USA*, 1996, 93: 5131–5135
- 39 Fang M, Andersson L. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. *Proc Biol Sci*, 2006, 273: 1803–1810
- 40 Joy D A, Feng X, Mu J, et al. Early origin and recent expansion of *Plasmodium falciparum*. *Science*, 2003, 300: 318–321
- 41 Rogers A R, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol*, 1992, 9: 552–569
- 42 Fu Y X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 1997, 147: 915–925
- 43 Bilgin R, Karatas A, Coraman E, et al. Regionally and climatically restricted patterns of distribution of genetic diversity in a migratory bat species, *Miniopterus schreibersii* (Chiroptera: Vespertilionidae). *BMC Evol Biol*, 2008, 8: 209
- 44 Robertson D L, Sharp P M, McCutchan F E, et al. Recombination in HIV-1. *Nature*, 1995, 374: 124
- 45 Sanderson M J, Doyle J J. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. *Syst Biol*, 1992, 41: 4–17
- 46 Zhou J, Bowler L D, Spratt B G. Interspecies recombination, and phylogenetic distortions, within the glutamine synthetase and shikimate dehydrogenase genes of *Neisseria meningitidis* and commensal *Neisseria* species. *Mol Microbiol*, 1997, 23: 799–812
- 47 Lau A N, Peng L, Goto H, et al. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. *Mol Biol Evol*, 2009, 26: 199–208
- 48 Klein J, Sato A, Nagl S, et al. Molecular trans-species polymorphism. *Ann Rev Ecol System*, 1998, 29: 1–21
- 49 Gleason C. The Biography of Silk. St. Catharines: Crabtree Publishing Company, 2006
- 50 Robinson R. Relaxing the clock brings time back into phylogenetics. *Plos Biol*, 2006, 4: e106
- 51 Gupta A K. Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Curr Sci*, 2004, 87: 54–59
- 52 Maekawa H, Takada N, Mikitani K, et al. Nucleolus organizers in the wild silkworm *Bombyx mandarina* and the domesticated silkworm *B. mori*. *Chromosoma*, 1988, 96: 263–269
- 53 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

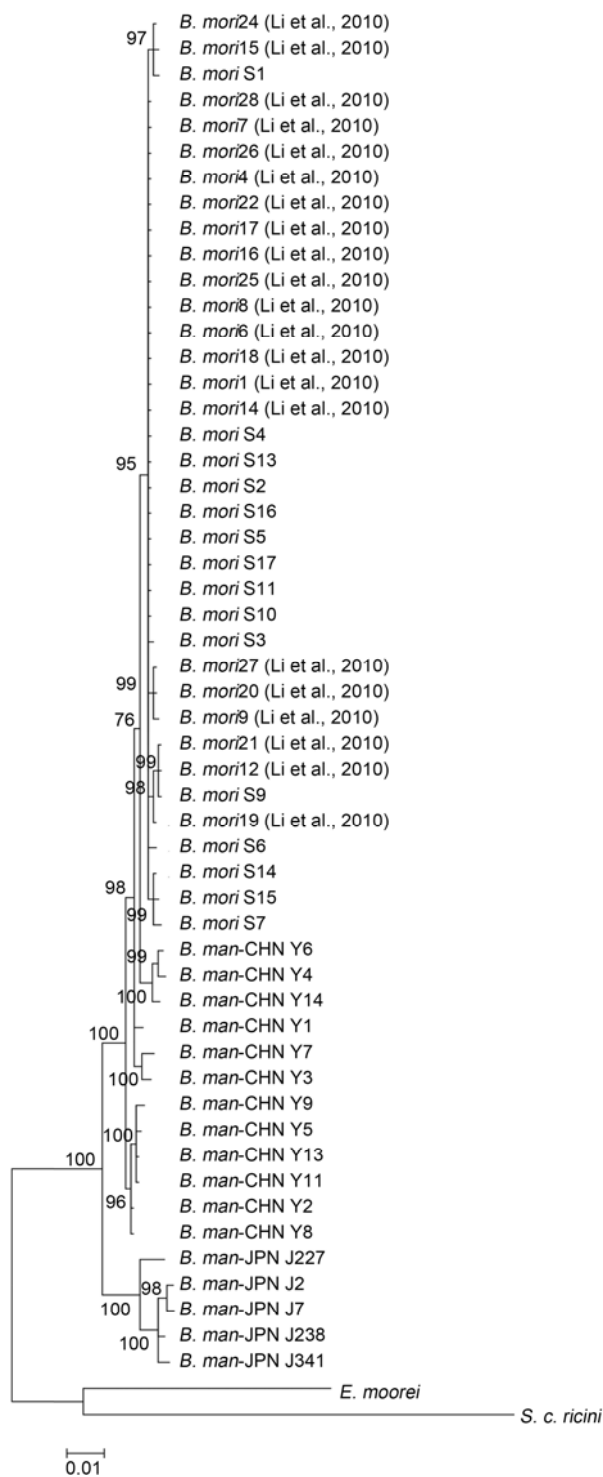
- 54 Santos C, Montiel R, Sierra B, et al. Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). *Mol Biol Evol*, 2005, 22: 1490–1505
- 55 Lambert D, Ritchie P, Millar C, et al. Rates of evolution in ancient DNA from *Adelie penguins*. *Science*, 2002, 295: 2270–2273
- 56 Edwards C J, Bollongino R, Scheu A, et al. Mitochondrial DNA analysis shows a Near Eastern neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc Biol Sci*, 2007, 274: 1377–1385
- 57 Beckenbach A T, Wei Y W, Liu H. Relationships in the *Drosophila obscura* species group, inferred from mitochondrial cytochrome oxidase II sequences. *Mol Biol Evol*, 1993, 10: 619–634
- 58 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
- 59 Crozier R, Crozier Y, Mackinlay A. The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates. *Mol Biol Evol*, 1989, 6: 399–411
- 60 Ho S Y, Phillips M J, Cooper, et al. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol Biol Evol*, 2005, 22: 1561–1568
- 61 Ho S Y, Larson G. Molecular clocks: when times are a-changin. *Trends Genet*, 2006, 22: 79–83
- 62 Marko P B. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Mol Biol Evol*, 2002, 19: 2005–2021
- 63 Benton M J, Ayala F J. Dating the tree of life. *Science*, 2003, 300: 1698–1700
- 64 Ho S Y, Phillips M J. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst Biol*, 2009, 58: 367–380
- 65 Luikart G, Gellie L, Excoffier L, et al. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci USA*, 2001, 98: 5927–5932
- 66 Lo C. Environmental impact on the development of agricultural technology in China: the case of the dike-pond ('jitang') system of integrated agriculture-aquaculture in the Zhejiang Delta of China. *Agr Ecosyst Environ*, 1996, 60: 183–195
- 67 Bray F. The case of China. *Education about Asia*, 2004, 9: 14–20

附 录



附图 1 基于 *TH* 序列的 BI 系统发生树

进化枝左上侧值: ML 树的 bootstrap 值; 右上侧值: MP 树的 bootstrap 值; 下侧值: BI 树的后验概率; *B. man*-CHN: 中国野桑蚕; *B. man*-JPN: 日本野桑蚕



附图 2 基于串联线粒体序列的 ML 系统发生树

进化枝上侧值: ML 树的 bootstrap 值; *B. man-CHN*: 中国野桑蚕; *B. man-JPN*: 日本野桑蚕; Li et al., 2010: 该序列从 Li et al., 2010 文章内获得