Phylogeny and evolutionary history of the silkworm

SUN Wei[1], YU HongSong[1], SHEN YiHong[1], BANNO Yutaka[2], XIANG ZhongHuai[1] & ZHANG Ze[1,3*]

[1]*Institute of Sericulture and Systems Biology, Southwest University, Chongqing 400715, China;*
[2]*Institute of Genetic Resources, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 812-8581, Japan;*
[3]*Institute of Agricultural and Life Sciences, Chongqing University, Chongqing 400044, China*

The silkworm, *Bombyx mori*, played an important role in the old Silk Road that connected ancient Asia and Europe. However, to date, there have been few studies of the origins and domestication of this species using molecular methods. In this study, DNA sequences of mitochondrial and nuclear loci were used to infer the phylogeny and evolutionary history of the domesticated silkworm and its relatives. All of the phylogenetic analyses indicated a close relationship between the domesticated silkworm and the Chinese wild silkworm. Domestication was estimated to have occurred about 4100 years ago (ya), and the radiation of the different geographic strains of *B. mori* about 2000 ya. The Chinese wild silkworm and the Japanese wild silkworm split about 23600 ya. These estimates are in good agreement with the fossil evidence and historical records. In addition, we show that the domesticated silkworm experienced a population expansion around 1000 ya. The divergence times and the population dynamics of silkworms presented in this study will be useful for studies of lepidopteran phylogenetics, in the genetic analysis of domestic animals, and for understanding the spread of human civilizations.

silkworm, domestication, divergence time, population dynamic, population expansion

The domesticated silkworm is the foundation of the silk industry in many developing countries, including China, India, and Brazil. More than 1000 inbred lines of the domesticated silkworm are maintained worldwide. It is generally believed that *Bombyx mori* was domesticated from an ancestor of the extant wild silkworm, *Bombyx mandarina* [1−4]. Currently, two types of wild silkworms exist; one occurs in China and in far eastern Russia (chromosome number 2*n*=56, as in *B. mori*), and the other in Japan and southern Korea (2*n*=54) [1,5]. They are usually referred to as Chinese *B. mandarina* and Japanese *B. mandarina*, respectively. Archeological evidence and cytogenetic information suggested that *B. mori* was first domesticated from *B. mandarina* in China and this is supported by molecular

phylogenetic studies [4,6–9]. Silkworm-rearing technology then spread to Korea and Japan around 200 BCE and was later introduced to central Asia and Europe through the Silk Road [10]. During this time, the exchange of silk products and silkworm-rearing technologies played an important role in the trade of goods and in the interaction of cultures among different countries. Hence, to some extent, the history of the spread of the silkworm reflects the history of commercial and cultural communication between ancient Asia and Europe. Although abundant historical records exist, molecular evidence supporting these records is still lacking. A recent study used the substitution rate of the mitochondrial *COI* and *COII* genes of another lepidopteran (Papilionidae) to estimate the divergence time of silkworms and concluded that domestication occurred about 1.08−1.41 million years ago (Ma) [11]. Another study, which used the

---

*Corresponding author (email: zezhang@swu.edu.cn)

substitution rate of the mitochondrial gene *Nd5* from a col-eopteran insect (Carabidae), estimated that *B. mori* split from wild silkworms about 7.1 Ma [12]. Both estimates are much older than indicated by the fossil evidence [4]. Therefore, the time of divergence of silkworms is still to be determined.

The genetic structures of domestic animals have been influenced by human activities and the migration of people. More than 1000 inbred lines of *B. mori* are maintained throughout the world and multiple phenotypes are present. However, previous analyses based on whole nuclear or mitochondrial genomes have concluded that the population structure of the domestic silkworm has remained stable since domestication [8,9]. Thus, further study of the population structure of *B. mori* is needed.

The aim of this study was to reveal the evolutionary history of the domesticated silkworm. Phylogenetic analyses of both mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) sequences were carried out. Our data again support the contention that the domesticated silkworm, *B. mori*, is closer to Chinese *B. mandarina* than to Japanese *B. mandarina*. Bayesian and maximum likelihood approaches under the relaxing molecular clock model were used to estimate the divergence time among the three silkworms. The results are consistent with the fossil evidence and historical records. Our results also revealed that *B. mori* experienced a population expansion about 1000 ya. This study, therefore, provides new insights into the evolutionary history of the silkworm.

1　Materials and methods

1.1　Taxon sampling and DNA extraction

Sixteen inbred *B. mori* lines representing four major geographic lines (Chinese, Japanese, European, and Tropical) were obtained from the silkworm resource pool at the Institute of Sericulture and Systems Biology, Southwest University, China. Thirteen samples of Chinese *B. mandarina* and five samples of Japanese *B. mandarina* were collected from different geographic regions in China and Japan. *Ernolatia moorei*, belonging to a sister genus of *B. mori* and, another species *Samia cynthia ricini* from the family Saturniidae, were used as outgroups in the phylogenetic analysis (Table 1).

Genomic DNA was extracted from individual pupae or moths using a standard phenol-chloroform extraction method [13].

1.2　Loci studied, amplification and sequencing

Two nuclear loci (the period gene, *period* and the tyrosine hydroxylase gene, *TH*) were chosen. Amino acid substitutions in period, especially in the PAS domain, occur much faster than in other protein-coding nuclear genes in Lepi-doptera (e.g., 4.9 times faster than in dopa decarboxylase and 44 times faster than in elongation factor-1α) [14]. Thus, we considered that this gene might exhibit a rapid evolutionary rate, which could distinguish it from closely related species. A recent study suggested that the *TH* gene has experienced artificial selection during domestication of the silkworm [15,16]. Therefore, *TH* was selected for the phylogenetic analyses of silkworms. Two classical mtDNA markers (*COI*, *CR*) were also used in this study.

Primers of four loci were designed based on the silkworm genome sequence and the complete mitochondrial genome sequence (Accession No. AY048187). The primers of the *TH* gene in *S. c. ricini* were based on the cDNA sequence published in a previous study [17]. For the *E. moorei TH* gene, we downloaded a number of lepidopteran *TH* coding sequences and identified the conserved region, from which we designed the primers. The amplification primers and sequences of the four loci are shown in Table 2. PCR reactions were performed to amplify the targeted sequences. The PCR products were subcloned into PMD-18-T vectors or PMD-19-T vectors (Takara Bio Inc., Dalian, China). Two or three positive clones were randomly selected and sequenced from each individual to avoid false polymorphism sites that may result from Taq polymerase errors [18].

1.3　Sequence alignment and analysis

The sequences were aligned using MUSCLE 3.6 [19] and corrected by eye. The sequences of the two mtDNA loci were combined to form a concatenated dataset, excluding samples that did not contain both complete sequences. The same method was used to construct a concatenated sequence dataset for the two nDNA loci.

1.4　Phylogeny reconstruction

Phylogenetic analyses were performed by three different methods: maximum likelihood (ML), maximum parsimony (MP), and Bayesian inference (BI). All trees of the mtDNA and *TH* genes were rooted by corresponding homologous sequences in *E. moorei*, a sister genus to *B. mori*, and/or *S. c. ricini*, which is another member of the Bombycoidea. The combinabilities of nuclear genes and mitochondrial loci were tested using the ILD test in PAUP [20,21]. The best fitting nucleotide substitution models for BI and ML tree reconstructions were searched by Modeltest Server 1.0 [22] under the Akaike Information Criterion (AIC), for each gene and the concatenated datasets.

The heuristic search option in PAUP* 4.0b10 [21] was employed for MP tree reconstructions, using 100 random sequence addition replicates with tree bisection reconnection (TBR) branch swapping [21]. Gaps were treated as missing data. The ML method was performed using PhyML 3.0 software [23]. The BioNJ tree was selected as the start-

Table 1 Summary of specimens used in this study

Species	Symbol	Name of strain	Source	Voltine[a]
B. mori	*B. mori* S1	Zhugui	Zhejiang, China	CV1
	B. mori S2	Huangbo	China	CV1
	B. mori S3	Sichuansanmian	Sichuan, China	CV1
	B. mori S4	Qiansanmian	Guizhou, China	CV1
	B. mori S5	Yindusanmian	India	T
	B. mori S6	Chishu	Japan	JV1
	B. mori S7	Ou18	Europe	EV1
	B. mori S8	Yi16	Europe	EV1
	B. mori S9	J115	Japan	JV2
	B. mori S10	872	Japan	JV2
	B. mori S11	DazaoN	China	CVd
	B. mori S13	WuG	Ukraine	EV2
	B. mori S14	WuF	Ukraine	EV2
	B. mori S15	Cambodia	Cambodia	T
	B. mori S16	BaixiaB	China	CV3
	B. mori S17	C108N	China	CV2
Chinese *B. mandarina*	*B. man*-CHN Y1	Huzhou	Huzhou, Zhejiang	Wild
	B. man-CHN Y2	Nanchong	Nanchong, Sichuan	Wild
	B. man-CHN Y3	Ankang	Ankang, Shanxi	Wild
	B. man-CHN Y4	Yichang	Yichang, Hubei	Wild
	B. man-CHN Y5	Luzhou	Luzhou, Sichuan	Wild
	B. man-CHN Y6	Lixian	Lixian, Hunan	Wild
	B. man-CHN Y7	Suzhou	Suzhou, JiangSu	Wild
	B. man-CHN Y8	Qingmuguan	Qingmuguan, Chongqing	Wild
	B. man-CHN Y9	Rongchang	Rongchang, Chongqing	Wild
	B. man-CHN Y11	Ziyang	Ziyang, Sichuan	Wild
	B. man-CHN Y12	Pengshan	Pengshan, Sichuan	Wild
	B. man-CHN Y13	Hongya	Hongya, Sichuan	Wild
	B. man-CHN Y14	Xiangyun	Xiangyun, Yunnan	Wild
Japanese *B. mandarina*	*B. man*-JPN J2	Higashi-Ku, Fukuoka	Fukuoka, Japan	Wild
	B. man-JPN J7	Fukuoka	Fukuoka, Japan	Wild
	B. man-JPN J227	Fukushima	Fukushima, Japan	Wild
	B. man-JPN J238	Shizuoka	Shizuoka, Japan	Wild
	B. man-JPN J341	Hokkaido	Hokkaido, Japan	Wild
E. moorei	*E. moorei*	Southeast Asia	Southeast Asia	Outgroup
S. c. ricini	*S. c. ricini*	Guangxi	Guangxi, China	Outgroup

a) CV1, Chinese univoltine; CV2, Chinese bivoltine; CV3, Chinese trivoltine; CVd, Chinese multi-voltine; JV1, Japanese univoltine; JV2, Japanese bivoltine; EV1, European univoltine; EV2, European bivoltine; T, Tropical; Wild, *B. mandarina*.

Table 2 Primer pair, sequence length and fitted substitution model for mitochondrial and nuclear DNA loci

DNA Type	Locus	Primer pair (5′→3′ primer sequence)	Length (bp)[a]	Model
Mitochondrial DNA	*COI*	F: GCCCTAAACAACCAAAAG R: GGCGACAAATCATAAAGATA	718	GTR+I
	CR	F: GCAACTGCTGGCACAAAT R: TGAGGTATGAGCCCAAAAGC	972	TrN+G
Concatenated mtDNA			1690	GTR+I
Nuclear DNA	*period*	F: ATAAACGAAGACCATTCGGT R: AATCCTTTCGTGATGCGTCA	1109	HKY+G
	TH	BmF[b]: AAGTGGTAAGGGTTGAGGGT BmR[b]: CGAAGGTTTTGTCTTGCTGT ErmF[c]: TGCATTGGTCTGTTTGCAAT ErmR[c]:G AACGACGGCGAGATGGGT SamF[d]: CATCAGATTCCGTCATTA SamR[d]: GCACCCAAATAAGTTCTA	969	TVM+I

a) Length of aligned sequences. b) The primer for amplifying the *BmTH* gene of *B. mori* and *B. mandarina*. c) The primer for amplifying the *TH* gene of *E. moorei*. d) The primer for amplifying the *TH* gene of *S. c. ricini*.

ing tree in each case. One thousand replicates were resampled to assess support for nodes in the MP and ML trees.

Bayesian inferences were performed using MrBayes v3.1.2 [24]. Four Markov chains on the data were run for 10^7 generations with one cold and three heated chains each, sampling once every 100 trees. To determine the burn-in, the online program AWTY [25] was used to plot the cumulative posterior split probabilities from the Markov chain Monte Carlo (MCMC) runs.

A haplotype network was reconstructed using the median-joining model implemented in the Network 4.516 program [26].

1.5   Detecting recombination

To detect potential recombinant individuals and the possible recombination breakpoints on the *period* gene, we employed the non-parametric RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA and SISTER SCAN recombination detection methods, as implemented in RDP3 [27]. Default program settings were used for all methods and a Bonferroni corrected P-value cutoff of 0.05 was used to minimize the probability of falsely inferring recombination.

1.6   Divergence time estimation

Different methods implemented in two software packages, BEAST v1.5.4 and r8s v1.7 [28,29], were used to estimate the time of divergence of silkworms, based on the concatenated mtDNA sequences. Three different parametric models of demographic history are implemented in BEAST: constant population size, exponential population growth, and expansion population growth. Two independent runs of 10^7 generations were performed for each BEAST setup. The first 25% of generations were discarded as burn-in. Convergence of parameters was checked and the mean and 95% highest posterior density interval (HPD) for the divergence time was calculated using Tracer v.1.5. BEAST needs one or more known calibrated points. One half of a fossilized silkworm cocoon shell was found in the prehistoric remains of the Xiyin village, which is believed to have existed 5000 ya, based on C-14 dating [30]. In addition, silk and silk products were found in the prehistoric remains of the late Neolithic Liangzhu village, which was also present 5000 ya [31]. These fossil records indicate that the silkworm was domesticated about 5000 years ago. A generally accepted study showed that agriculture originated during the Neolithic period [30]. Therefore, a uniform prior with broad upper and lower bounds of 0.004 and 0.011 Ma was fixed to calibrate the estimate. In addition, a recent study estimated that the divergence of Japanese wild silkworms and Chinese wild silkworm occurred about 1.11–1.45 Ma [11]. This time estimate was used as another calibrated point.

Two methods, nonparametric rate smoothing (NPRS) and semi-parametric penalized likelihood (PL) [32,33] implemented in r8s, were used to estimate the divergence time among silkworms. Both methods used the mtDNA ML tree as the input tree (without outgroups). We also constrained the dates to be the same as used in the BEAST software (i.e., divergence of the domesticated silkworm and Chinese wild silkworm 0.004–0.011 Ma; split time of wild silkworms 1.11–1.45 Ma). The Powell algorithm was applied to optimize branch lengths in the NPRS method, and the truncated Newton algorithm was used in the PL method. The 95% confidence interval (CI) of the estimated data was calculated by the bootstrap method with the r8s-bootstrap kit [34]. This kit also used the SEQBOOT program in the Phylip package v3.6 (http://www.bioinformatics.uthscsa.edu/www/phylip/).

1.7   Historical demography

The past population dynamics of the domesticated silkworm and Chinese wild silkworm was examined using two methods in the software Arlequin v3 [35], based on the concatenated *COI* and *CR* sequences. The Japanese wild silkworm was excluded in this analysis because of its small population size. First, the pairwise mismatch distributions among individuals were calculated with 1000 bootstrap replications. Then, Fu's F_s test was used to determine the excess of rare alleles that could indicate a recent expansion [36]. To obtain more information concerning the history of population dynamics of the domesticated silkworm, Bayesian skyline plots (BSP) implemented in BEAST v1.5.4 were also performed [28].

2   Results

2.1   Sequence data characteristics

Sixteen *B. mori*, thirteen Chinese *B. mandarina* and five Japanese *B. mandarina* samples were selected for this study (Table 1). The two classic mtDNA markers (*COI* and *CR*) and the two single-copy nuclear loci (*period* and *TH*) were sequenced in almost all of the domesticated and wild silkworms (the *CR* regions of one domesticated silkworm *B. mori* S8 and one Chinese wild silkworm *B. man*-CHN were not amplified; and the *period* genes of one domesticated silkworm *B. mori* S4, one Japanese wild silkworm *B. man*-JPN J341, and one outgroup *S. c. ricini*; were not amplified). Multiple-sequence alignment showed that the domesticated silkworm has a higher sequence similarity with the Chinese wild silkworm than with the Japanese wild silkworm. Most of the differences represented divergence between the Japanese wild silkworm and Chinese wild silkworm/*B. mori* (Table 1). In addition, the Japanese wild silkworm possessed many unique indels at mtDNA and nDNA loci, e.g., a 252 bp inserted fragment in *CR*, and two unique insertions (285 and 45 bp) in *period*.

All sequences used in this study were deposited in Gen-Bank (GenBank accession Nos. *COI*: GQ423211–GQ423244, GU360733, HQ132262; *CR*: GQ423245–GQ423276, GU360735, HQ132261; *period*: HQ132228–HQ132259, HQ132227; *TH*: GQ423342–GQ423346, GQ415508–GQ415537, GU360734, HQ132260).

2.2 Phylogenetic analysis

The incongruence length difference test, also known as the partition homogeneity test, was used to determine whether the sequences from the two nuclear genes, two mitochondrial loci or these four loci could be concatenated [20]. The results showed no conflict within the mtDNA (*COI-CR*, P =

0.349), but conflict within the nDNA (*period-TH*, $P=0.001$) and between the nuclear and mitochondrial loci (nDNA-mtDNA, $P = 0.001$), suggesting that it is unreasonable to combine the two nuclear genes and all four loci. Thus, we performed phylogenetic analyses for the concatenated mtDNA, *period* and *TH* sequences.

The phylogenetic trees based on the concatenated mtDNA, *period* and *TH* showed similar topologies. Therefore, only the concatenated mtDNA tree (Figure 1) and the *period* tree (Figure 2) are shown in the main text. All trees showed two main and distantly related clades. The first clade, designated as BJ, was a monophyletic group comprising only the Japanese wild silkworm with high bootstrap supports and Bayesian posterior probabilities in all three
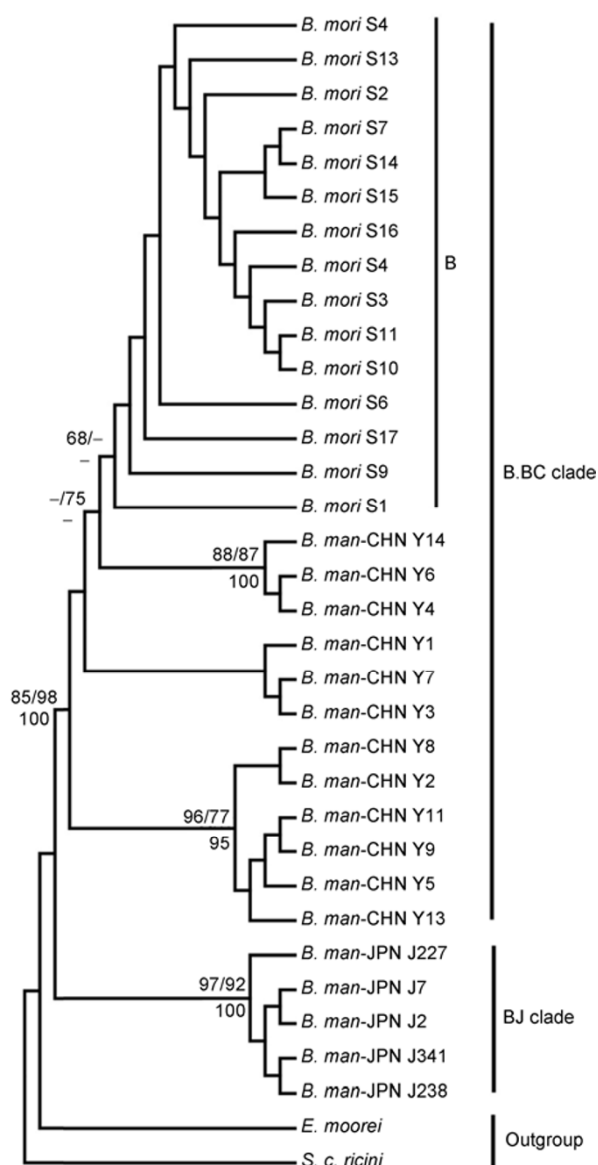

Figure 1   MP Phylogenetic tree reconstruction based on the concatenated mtDNA. Support values above the branches represent ML bootstrap (left) and MP bootstrap values and the values below the branches represent Bayesian posterior probabilities. "–" indicates bootstrap values or BPP less than 50. *B. man*-CHN is the Chinese wild silkworm and *B. man*-JPN is the Japanese wild silkworm. *E. moorei* refers to *Ernolatia moorei* and *S. c. ricini* refers to *Samia cynthia ricini*.
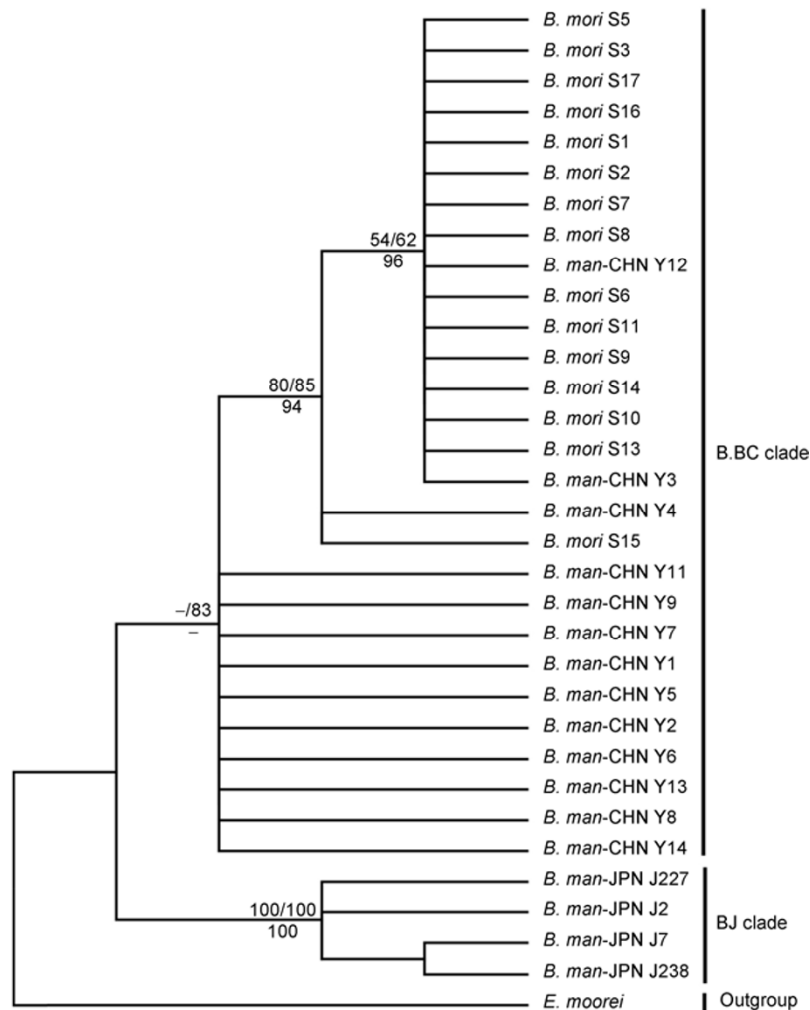
Figure 2   MP Phylogenetic tree reconstruction based on the *period* gene. Support values above the branches represent ML bootstrap (left) and MP bootstrap values and the values below the branches represent Bayesian posterior probabilities. "–" indicates bootstrap values or BPP less than 50. *B. man*-CHN is the Chinese wild silkworm; *B. man*-JPN is the Japanese wild silkworm. *E. moorei* refers to *Ernolatia moorei* and *S. c. ricini* refers to *Samia cynthia ricini*.

phylogenetic analyses. The second main clade, designated as B.BC, consisted of *B. mori* and Chinese *B. mandarina*, indicating a closer relationship between the domesticated silkworm and the Chinese wild silkworm. However, the B.BC-clade topology changed a little when different locus sequences were used to reconstruct the phylogenetic tree. In the *period* tree, there was a mixed clade containing all of the domesticated silkworm and three Chinese wild silkworm samples (*B. man*-CHN Y3, *B. man*-CHN Y4 and *B. man*-CHN Y12) (Figure 2). In contrast, in the mtDNA and *TH* trees, all of the domesticated silkworm samples formed a single group (B clade), which was clustered with the Chinese wild silkworm (Figure 1; Appendix Figure 1 in the electronic version). Furthermore, when we added the *CR* and *COI* sequences from other domesticated silkworm lines sequenced in [8] to reconstruct the phylogenetic tree, we obtained a similar result (Figure 1; Appendix Figure 2 in the electronic version). In summary, our analysis suggested that the domesticated silkworm has a closer relationship with the

Chinese wild silkworm than with the Japanese wild silkworm, supporting the notion that the domesticated silkworm was directly domesticated from the Chinese wild silkworm.

2.3   Divergence time estimates

Two different methods, the Bayesian Markov chain Monte Carlo (MCMC) method (in BEAST) and the maximum likelihood method (in r8s), were used to estimate the divergence time among silkworms, based on the concatenated mtDNA sequences.

Based on historical and fossil evidence, the split date of the domesticated silkworm and the Chinese wild silkworm (0.004–0.011 Ma) was used as the first calibrated point for each demographic model (Table 3, gray background). For BEAST, three parametric models of demographic history (constant population size, exponential population growth, and expansion population growth) were used to estimate the divergence time. The predictions obtained from these three

Table 3   Mean node ages and confidence intervals attained using different analytical approaches[a]

Software loci	Model/method	Age estimates (ya) (95%HPD)		
		B. man-JPN vs. *B. man*-CHN	*B. man*-CHN vs. *B. mori*	*B. mori* vs. *B. mori*
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) * represents the calibrated point; gray background represents the results calculated using the first calibrated point (0.004–0.011 Ma); white background represents the results calculated using the second calibrated point (1.11–1.45 Ma).

demographic models were very similar. Therefore, only the chronology of the divergence time obtained from the constant population size model is shown in Figure 3. The most recent common ancestor of the Chinese wild silkworm and the domesticated silkworm was estimated to have existed about 4100 ya. The estimate for the time of separation of the Chinese wild silkworm and the Japanese wild silkworm was 23600 ya, with lower and upper bounds of the 95% highest posterior density interval (95% HPD) of 5200 and 50200 ya. This time is much earlier than the time of the domestication event. The date of divergence of wild silkworms was also confirmed by analyses under the exponential population growth model (14700 ya; 95% HPD 4000–28300 ya) and the expansion population growth model (20900 ya; 95% HPD 4100–39100 ya).

The estimate for the time of radiation of the different geographic lines of *B. mori* under the constant population size model was 2000 ya (95% HPD 600–4300 ya), which was similar to values obtained by the other two models for population change (Table 3). The divergence time of silkworms was also estimated using the r8s v1.7 software. Two maximum likelihood dating methods, NPRS and PL, were used in these analyses [32,33]. Both estimates were similar to those obtained from BEAST (Table 3). The estimate for the time of radiation of the domesticated silkworm using NPRS was 1540 ya (95% CI 716–2811 ya), and using PL was 2891 ya (95% CI 2085–3605 ya). The estimate of the split time of Chinese and Japanese wild silkworms was 46373 ya (95% CI 15931–65743 ya) using NPRS, and 53798 ya (95% CI 21642–78438 ya) using PL (Table 3). Overall, the times of the radiation of the domesticated silkworm, and of the split between Chinese and Japanese wild silkworms, estimated by the two software packages (BEAST and r8s) were consistent with the fossil evidence.

A recent study showed that the divergence of Japanese and Chinese wild silkworms occurred 1.11–1.45 Ma [11].

Thus, we used this date as the second calibrated point to calculate the divergence time (Table 3, white background). For BEAST software, the three different demographic models were also utilized in this analysis. This analysis indicated that the domestication event occurred 286800 ya (95% HPD 137900–473900 ya) and that the radiation of *B. mori* took place 117100 ya (95% HPD 46800–217200 ya). We also used the r8s software with the second calibration date (Table 3). By this method, the split time between the domesticated silkworm and the Chinese wild silkworm was calculated as 750000 ya (95% CI 586919–818598) using NPRS, and 410000 ya (95% CI 331947–544417) using PL. The corresponding time of spread of the domesticated silkworm was 471667 ya (95% CI 400933–542400) using NPRS and 400000 ya (95% CI 236644–659218) using PL. In each case, the time of domestication of silkworms appeared to be much earlier than that indicated by the fossil evidence and historical records [4,30,31]. Thus, it is unreasonable to use the divergence time of Japanese and Chinese wild silkworms cited in [11] as the calibration date.

2.4   Historical demography

The median-joining phylogenetic network was reconstructed based on the haplotype data of the concatenated mtDNA sequences (the sequences used in this analysis included the *CR-COI* regions sequenced in this study and the region sequenced in [8]) to generate a directed acyclic graph that models the evolutionary histories of the domesticated and wild silkworms (Figure 4A). The domesticated silkworm and the two wild silkworms formed three independent groups. As seen in the phylogenetic analyses, the Chinese wild silkworm had a close relationship with the domesticated silkworm. In contrast to the wild silkworm groups, the domesticated silkworm group exhibited a single core haplotype containing nine samples and a star-like pattern. Pre-
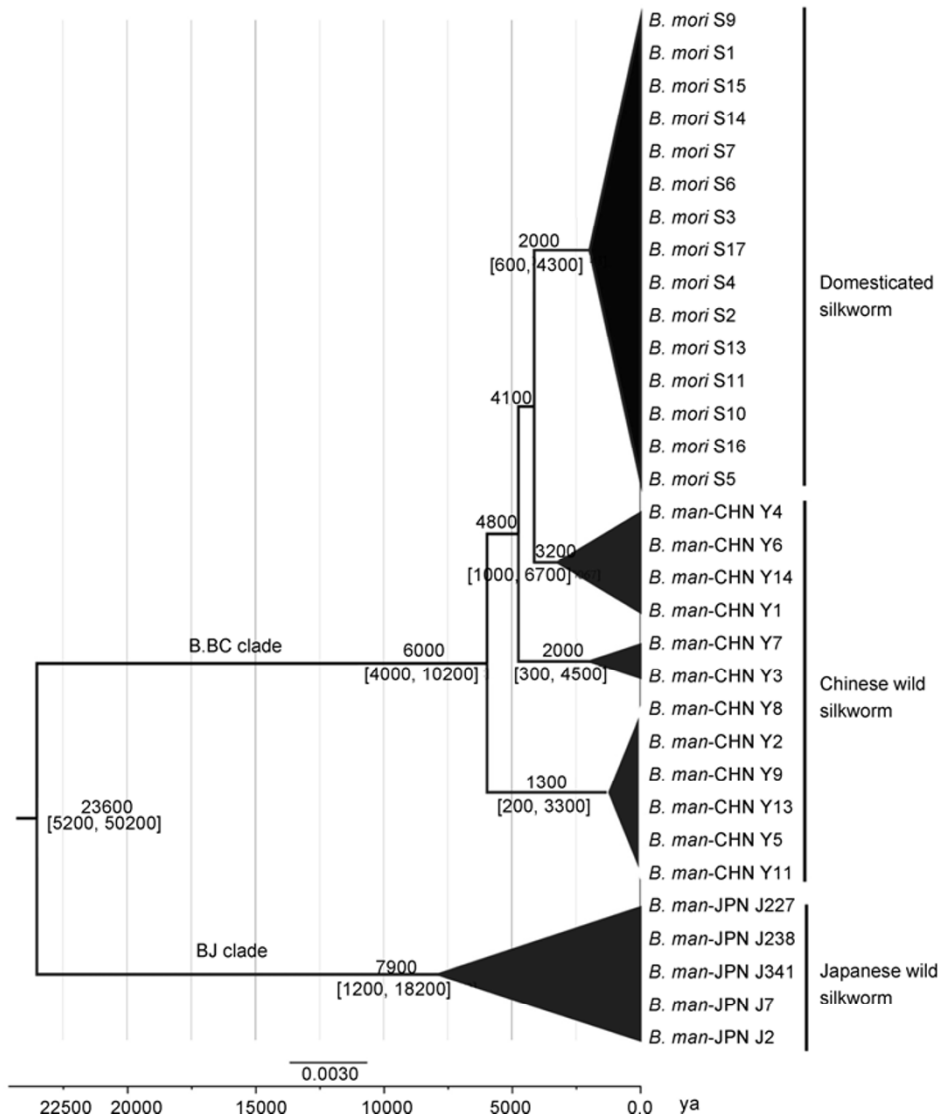
Figure 3   Chronology of divergence times among silkworms based on mtDNA, estimated using Bayesian methods implemented in BEAST. The value above the branch represents the divergence time and below branch is the 95% highest posterior density interval (HPD). *B. man*-CHN refers to the Chinese wild silkworm; *B. man*-JPN refers to the Japanese wild silkworm.

vious studies in domestic cattle and goats also revealed a star-like haplotype pattern, suggesting that these animals might have experienced population expansions [38,39]. The star-like pattern in our network analysis thus suggests that the domesticated silkworm might also have experienced a population expansion [26,40]. We also used the complete mtDNA sequences of the domesticated silkworm and the Chinese wild silkworm in [8] to reconstruct the haplotype network (Figure 4, inset). The relationship of haplotypes among the domesticated silkworm was very close. Because of high haplotype diversity (Hd=1.0), no core haplotype was detected although it also showed a star-like trend. These observations appear to be inconsistent with two previous studies that suggested the domesticated silkworm retained a stable population size in the past [8,9]. Therefore, more sophisticated analytical methods will be needed to further elu-

cidate the historical demography of silkworms. In addition, the *period* trees suggested some mixing between Chinese wild and domesticated silkworms (Figure 2). We used the haplotype network analysis to survey whether these silkworms have similar haplotypes. The analysis showed that three Chinese wild silkworm samples had similar haplotype patterns to that of the domesticated silkworm (Figure 4B).

The pairwise mismatch distribution and the value of Fu's F_s were calculated to measure the population structure of the domesticated silkworm and Chinese wild silkworm using Arlequin v3.0 [35]. The pairwise mismatch distribution is a frequency graph of pairwise differences between alleles. A smooth bell-shaped curve of the distribution indicates population expansion [41]. Fu's F_s test is very sensitive to an excess of recent mutations and is widely used to analyze population structure. A negative value of F_s indicates recent
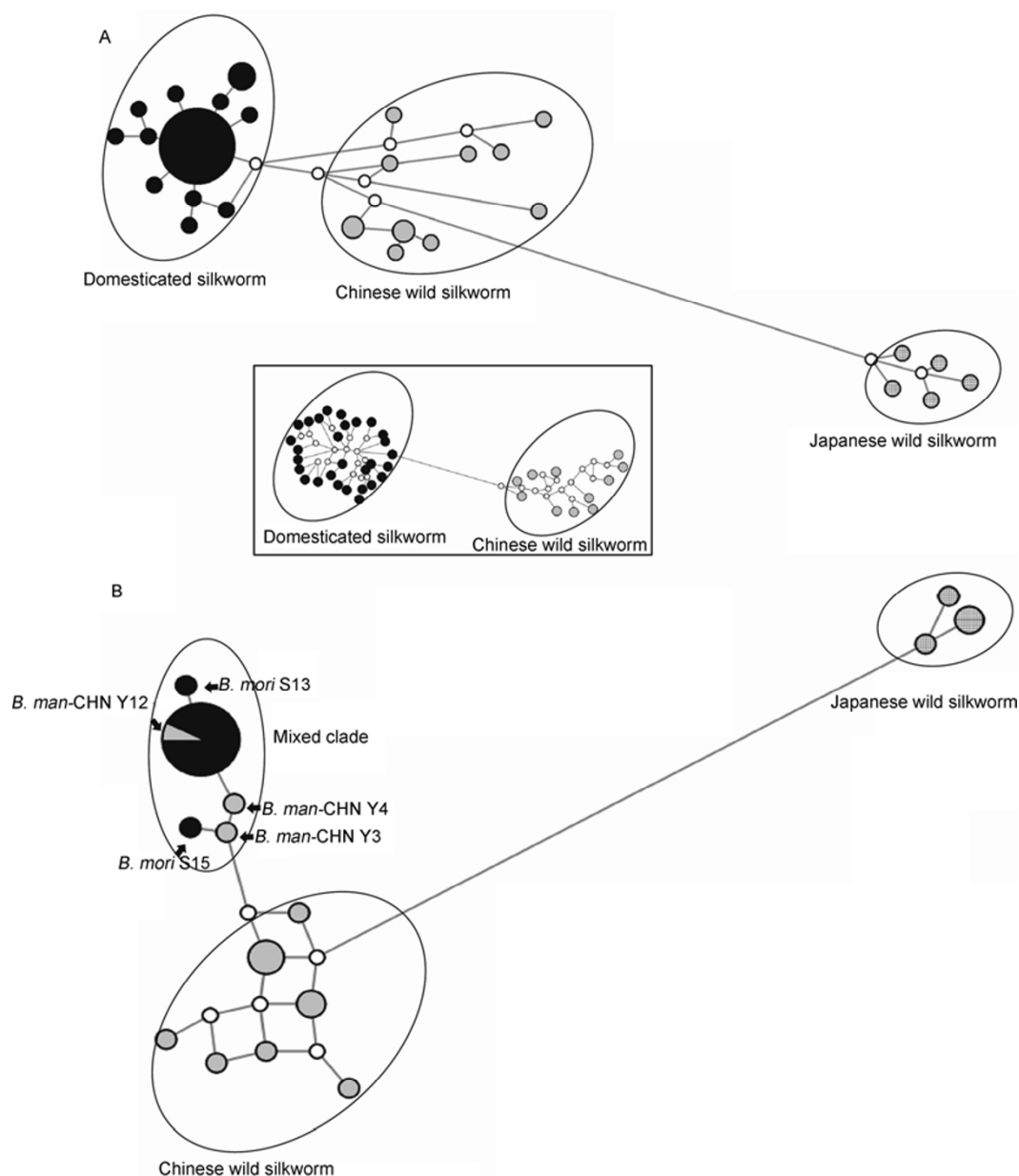
Figure 4   Parsimonious median-joining haplotype network analysis. A, The network based on the concatenated mtDNA sequences; the inset figure is the network based on the complete mtDNA sequences. B, The network based on the *period* gene. The sizes of the circles and the shaded segments are proportional to the haplotype frequencies in the datasets. Solid circles represent domesticated silkworms; gray circles represent Chinese wild silkworms; cross-hatched circles represent Japanese wild silkworms.

population growth [42]. The Japanese wild silkworm was excluded from this analysis because of its small population size (*N*=5). A single peak at a difference around 1 in the mismatch distribution analysis of the domesticated silkworm and the significant negative value of F_s (F_s=−3.769, *P*=0.003) showed that *B. mori* experienced population expansion whereas Chinese wild silkworm had no population expansion during its evolution (F_s=−2.252, *P*=0.110) (Figure 5A). The Bayesian skyline plots (BSP) with the MCMC sampling procedure also showed the population expansion of the domesticated silkworm and indicated that the population expansion probably occurred about 1000 ya (Figure 5B) [43].

We also measured the population structures of silkworms using the complete mitochondrial genome sequences of 29 *B. mori* strains and 11 Chinese *B. mandarina* samples [8]. Figure 5C shows a smooth curve with a single peak at around difference 15 in the mismatch distribution and a significant negative value of F_s (F_s=−19.131, *P*<0.001), consistent with the expansion of the domesticated silkworm population. The BSP analysis revealed that the expansion events probably started around 1000 ya (Figure 5D). The pairwise mismatch distribution using the complete mitochondrial genome data for 11 Chinese *B. mandarina* showed no population expansion for Chinese wild silkworm,
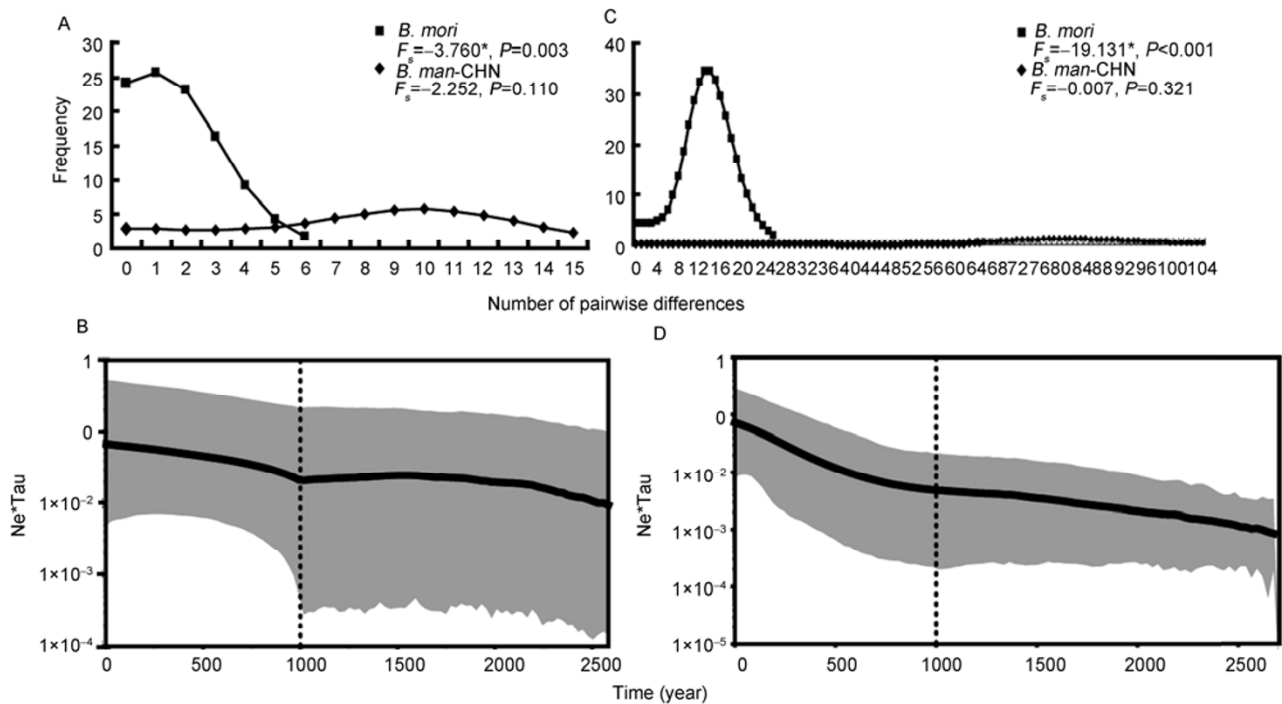
Figure 5   Pairwise mismatch distribution and Bayesian skyline plots in concatenated mtDNA (A and B) and the complete mitochondrial genome (C and D). In A and C, the lines with squares represent the domesticated silkworm; the lines with the diamonds represent the Chinese wild silkworm. In B and D, the gray region represents the 95% highest probability density interval. The dotted lines in B and D represent the potential age of the population expansion.

as shown above, and Fu's F_s test confirmed this ($F_s=-0.007$, $P=0.321$) (Figure 5C).

3   Discussion

3.1   Chinese *B. mandarina* and *B. mori* have a recent common ancestor

Nucleotide sequences of two mitochondrial loci and two nuclear loci obtained in the present study showed high similarities between *B. mori* and Chinese *B. mandarina*. In contrast, the sequences from Japanese *B. mandarina* have a number of unique characteristics. For example, Japanese *B. mandarina* has a 216 bp repetitive insertion in the mitochondrial *CR* region and two large unique insertions in the intron of the *period* gene. These observations are consistent with a previous finding that a transposon-like sequence was inserted in the *Arylphorin* gene of Japanese *B. mandarina* [5]. These observations all indicate that the molecular markers we selected had enough polymorphisms to distinguish the phylogenetic relationships among closely related species and that *B. mori* and Chinese *B. mandarina* are more closely related than are *B. mori* and Japanese *B. mandarina*.

The phylogenetic analyses based on the concatenated mtDNA sequences and two single nuclear genes all led to similar topologies. They all consistently revealed that Japanese *B. mandarina* first diverged from the other two silkworms (Figures 1 and 2). The Chinese wild silkworm and

the domesticated silkworm were then grouped together into a single clade, B.BC (Figures 1 and 2). The results are consistent with previous analyses which used the complete nuclear and mitochondrial genome sequences [8,9]. However, the topology of the B.BC clade changed a little when different locus sequences were used. For the concatenated mtDNA and *TH* sequences, the domesticated silkworm clade (B clade) first formed one clade and then grouped with the Chinese wild silkworm clade. However, the *period* trees suggested that these two silkworms were mixed together (Figure 2), and the haplotype network analyses also showed that three Chinese wild silkworm samples have similar haplotype patterns to that of the domesticated silkworm (Figure 4). Many previous analyses have indicated that recombination has a significant impact on phylogenetic accuracy [44–46]. To examine whether individual recombination events occurred within the *period* gene, a set of six recombination detection methods implemented in RDP3 [27] were used to detect evidence of recombination. These analyses showed no evidence of recombination events in the *period* gene. Similar difficulties were observed in the genetic analysis of the domesticated and wild horse, based on autosomal loci [47]. The ambiguous phylogenetic relationships between the domesticated animals and their wild ancestors, especially for the nuclear genes, may be explained by introgression or shared ancient polymorphisms [47]. Another study pointed out that ancestral polymorphism may influence phylogenetic accuracy, particularly for two closely related species with a divergence time of less than 1–2

million years [48]. In fact, our results suggested that *B. mori* and *B. mandarina* split only about 4100 years ago. We speculate that there may have been ancestral polymorphism at the *period* locus in the ancient wild silkworm population and that the silkworms captured for domestication included one of these haplotypes. Consequently, the phylogenetic tree based on the *period* locus results in a mixed clade containing the domesticated and wild silkworms.

3.2   The evolutionary history of silkworms

Evidence form archeology and cytogenetics suggested that *B. mori* was first domesticated in China from *B. mandarina* at least about 4600 ya [4]. Around 200 BCE, silkworm-rearing technology spread from China to Korea as Chinese people moved into that area, and then spread to Japan [49]. As the Silk Road was founded by Chinese traders from the Han Dynasty (206 BCE–220 BCE), ancient Asia had many connections with ancient Europe. The Silk Road enhanced the goods trade among many countries and the communication of different cultures. During this time, the silkworm-rearing technology and silk products were also introduced to Central Asia and Europe from China via the Silk Road [10]. Thus, to some extent, the history of the spread of silkworm from China to other countries is a representation of the history of trade and culture between ancient Asia and Europe. Although there are many historical records from this period, no previous molecular evidence was available to support them. As shown above, a previous study estimated that the divergence between Japanese and Chinese wild silkworms occurred 1.11–1.45 Ma and that their domestication occurred 1.08–1.41 Ma. The latter estimate is quite unlikely because agricultural history is much more recent [11,30]. Another study concluded that *B. mori* separated from Japanese *B. mandarina* about 7.1 Ma [12]. The large difference between these two studies may result from differences in the calibration of molecular clock rates. The former used the substitution rate of the mitochondrial *COI* and *COII* genes of another lepidopteran (Papilionidae) and the latter used the substitution rate of another mitochondrial gene *Nd5* from a coleopteran insect (Carabidae). Both of these two insects are phylogenetically distant from Bombycidae. Furthermore, both studies used linear regression to estimate the divergence time assuming a constant substitution rate across every lineage [28]. However, this simple model of substitution rate may not be realistic; the evolutionary rate always varies both within and between lineages [50].

We used three different population growth methods under a relaxed molecular clock model in the Bayesian MCMC analyses and two parametric methods implemented in maximum likelihood analyses to estimate the divergence time of silkworms based on mtDNA sequences. All of these assumptions yielded similar results, although some ages calculated in the Bayesian MCMC analyses were a little larger than those obtained by maximum likelihood analyses

(Table 3, Figure 3). The time of the split between *B. mori* and Chinese *B. mandarina* was estimated to be about 4100 ya, which is consistent with the fossil data [4]. The domestication date of silkworms is later than that of other domesticated species, e.g., dogs and rice. Dogs were first domesticated for hunting wild animals and rice was domesticated for food, both of which were important for ancient human survival [51]. Since the silk is not an absolute necessity for human existence, domestication of silkworm occurred later than other animals. The most recent common ancestor of the four main geographic populations of *B. mori* existed about 2000 ya (95% HPD 600–4300 ya). This time is in accordance with the historical records that *B. mori* spread into Japan and other countries nearly 2000 ya [37,49].

We estimated that the Chinese wild silkworm and the Japanese wild silkworm diverged about 23600 ya (95% HPD 5200–50200 ya), which is a little earlier than estimates of the time that the islands of Japan were completely separated from the Asian continent, about 20000 ya [51]. As shown above, Chinese wild silkworms and Japanese wild silkworms have different chromosome numbers and several studies have shown that the northern part of the Korean Peninsula may be the boundary dividing the two geographic *B. mandarina* populations. Thus, to calculate the divergence time of them, it would be better to use the Korean population. However, to date, it has been difficult for us to collect wild silkworm from Korea. Two previous studies indicated that the wild silkworms from the Korea resembled those from Japan, e.g., the same chromosome number and the same retroposon-like insertion in the arylphorin gene [52,53]. Therefore, we consider that adding the Korean population may have little impact on the calculation of the divergence time between the two wild silkworms. We also estimated the divergence time based on the nuclear *TH* gene (Table 3). All of the divergence time estimates using the *TH* gene sequences were lower than those using mtDNA sequences. Because BEAST software uses the phylogenetic tree as the framework to test the evolutionary hypotheses, the different topologies of the mtDNA and *TH* trees, particularly of the clades within the B.BC clade, may have led to the different divergence time estimates in this study. Furthermore, a recent study showed that *TH* has experienced strong artificial selection during silkworm domestication; this process may reduce the DNA polymorphism and affect the divergence time estimation [16].

Based on the fossil record, we estimated the molecular substitution rate of *COI* and *CR* in silkworms (*Bombyx*) using the BEAST program. The estimate of the substitution rate is about 1.024×10^{-6} (95% HPD 0.358×10^{-6}–1.872×10^{-6}) per site per year. This rate is similar to that of some mammals (e.g., *Homo sapiens*: 0.164×10^{-6}–1.398×10^{-6} [54]; *Pygoscelis adelia*: 0.53×10^{-6}–1.43×10^{-6} [55]; *Bos primigenius*: 0.15×10^{-6}–1.30×10^{-6} [56]) but it is much faster than that in some other insects (e.g., *Drosophila obscura*:

2.0×10^{-8}–2.9×10^{-8} [57]; *Papilio*: 7.8×10^{-9}–10.2×10^{-9} [58]; *Apis mellifera*: 3.4×10^{-8}–4.9×10^{-8} [59]). Ho *et al.* [60] showed that rates estimated over the short-term (<1–2 Ma) are time-scale dependent. Several factors may result in the patterns in rate estimates, for example, sequence error, purifying selection, calibration points, mutational hot-spots, and saturation [60]. For domesticated species, the calibration points may be the most common reason to overestimate the substitution rate because of short time frame of domestication events [61,62]. A fossilized half silkworm cocoon shell dated at 5000 ya was found in prehistoric remains [30]. Here, we used this fossil record as the calibration point to calculate the divergence time and the molecular substitution rate of silkworms. However, the fossil record only represents the minimum age [63]. Thus, we also used a maximum age, the date of the origin of Chinese agriculture, to fix the calibration point [30]. The fossil record with a maximum age constraint may be useful to estimate the molecular substitution rate [64]. Nevertheless, the extreme substitution rate we estimated may have been elevated by the relatively short time since this calibration point. In addition, the mtDNA loci of the domesticated silkworm were under strong purifying selection (Tajima's $D=-1.94459$, $P<0.05$), which also has an effect on the time-dependence of molecular rate estimates [60]. Regardless, the substitution rate estimated in this study will be useful for phylogenetic and genetic analyses of *Bombyx* and related species.

More than 1000 inbred lines of domesticated silkworms are maintained worldwide. Most of them show multiple phenotypes, such as different color patterns or voltines. Previous studies suggested that human activity may influence the genetic structures of domesticated animals and generate many different animal phenotypes [39]. During domestication, the demographic populations of domesticated animals have experienced expansions, e.g., cattle, goats and pigs [38,65]. However, previous investigations suggested that the population size of *B. mori* has been constant during its domestication [8,9]. It is difficult to understand the different population patterns between the domesticated silkworms and other domesticated animals. Thus, further studies are required to analyze the population history of the domesticated silkworm. In this study, four independent methods, haplotype networks, pairwise mismatch distributions, F_{st} tests, and Bayesian skyline plots, were used to analyze the population dynamics of the domesticated silkworm. All analyses strongly indicated that the domesticated silkworm experienced a recent population expansion starting about 1000 ya, and that there was no expansion of the Chinese wild silkworm (Figure 5A and B). We also performed the same analysis using the complete mitochondrial genomes sequenced in [8] of 29 *B. mori* and 11 Chinese *B. mandarina*. A strong indication of a population expansion in the domesticated silkworm was detected, which contrasts with the previous study [8]. However, both our analysis and the previous analysis showed similar evolutionary patterns

for the Chinese wild silkworm (Figure 5C and D). As shown in Figure 5, the mismatch distribution plot and F_s values exhibited differences when we used concatenated mtDNA and 29 complete mitochondrial sequences, which may be due to more polymorphic information supplied by the complete mitochondrial sequences. Nevertheless, we have shown that the domesticated silkworm experienced a population expansion as observed in other domesticated animals [38,39,65]. It is unclear why this conclusion differs from previous analyses [8,9]. A possible reason is that the dynamic model is too simple and does not consider bottlenecks [9]. We are unable to explain the discrepancy between our conclusions and those of another earlier study [8] because the authors did not show details of their tests concerning population dynamics. These conflicts indicate that further investigations into the population dynamics of silkworms are needed. The genetic structures of other domestic animals have changed as they have been adapted to the interests of humans, e.g., color, quality of meat. Similarly, selection has also occurred in the domestication of silkworms related to the quality of silk, and consequent changes in genetic information would be expected. Thus, we suggest that the model of population expansion is more plausible for *B. mori*. The expansion of the domesticated population was estimated to have occurred about 1000 ya, which was around the time of the Chinese Song Dynasty (960 CE–1279 CE). Before the Song Dynasty, the political and economic center was in the north of China. As the capital of the Song Dynasty moved to the south of China, the people also migrated from north to south and brought with them advanced agricultural technology [66]. During this period, it is possible that the domesticated silkworm lines from the north might have changed some of their characters as they adapted to the warmer temperatures and other environmental conditions in the south. Hybridization between the domesticated silkworms from two places might also have changed their genetic structures. Agricultural technology greatly improved during the Song Dynasty, which is referred as a "Green Revolution" [67], and it may have facilitated improvements in the quality of the domesticated silkworm. Taken together, these observations suggest that the period of population expansion of the domesticated silkworm (1000 ya) coincided with the migration of people and agricultural development which occurred during the Song Dynasty.

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, Kllersj 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 Gu G D, Xu J L. On the hearths of sericulture in China (in Chinese). J Zhejiang Univ Nationalities, 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 timesare 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, Gielly 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