

RESEARCH ARTICLE

Open Access

Demographic history and gene flow during silkworm domestication

Shao-Yu Yang^{1†}, Min-Jin Han^{2†}, Li-Fang Kang², Zi-Wen Li¹, Yi-Hong Shen^{1*} and Ze Zhang²

Abstract

Background: Gene flow plays an important role in domestication history of domesticated species. However, little is known about the demographic history of domesticated silkworm involving gene flow with its wild relative.

Results: In this study, four model-based evolutionary scenarios to describe the demographic history of *B. mori* were hypothesized. Using Approximate Bayesian Computation method and DNA sequence data from 29 nuclear loci, we found that the gene flow at bottleneck model is the most likely scenario for silkworm domestication. The starting time of silkworm domestication was estimated to be approximate 7,500 years ago; the time of domestication termination was 3,984 years ago. Using coalescent simulation analysis, we also found that bi-directional gene flow occurred during silkworm domestication.

Conclusions: Estimates of silkworm domestication time are nearly consistent with the archeological evidence and our previous results. Importantly, we found that the bi-directional gene flow might occur during silkworm domestication. Our findings add a dimension to highlight the important role of gene flow in domestication of crops and animals.

Keywords: Coalescent simulation, Approximate Bayesian computation, Gene flow, *Bombyx mori*

Background

The last decade has seen the advances in the field of demographic history involving gene flow of domesticated species [1,2]. Initially, the studies of demographic history involving gene flow were mainly focused on crop plants [3]. The studies of domesticated animals, although they are crucial for human beings, are relatively lagged behind that of crop plants [4]. Until recently, some studies have investigated demography of domesticated animals, such as dog [5], chicken [6] and pig [7]. When considerable attention has been focused on these species, the same as domesticated species, few similar evolutionary studies have involved the domesticated silkworm, *Bombyx mori*.

The domesticated silkworm, *B. mori*, is an economically important insect, and widely distributed in many regions especially China, Korea, Japan and India [8]. In addition to its economic importance for farmers in these

countries, the domesticated silkworm is a model of lepidopteran insects in basic research because of its short life cycle, and adaptation to laboratory culture [8]. The draft genome sequences of the domesticated silkworm were completed in 2004 [9,10] and a new assembly of the genome sequence as well as the resequencing data were released in 2008 [11] and in 2009 [12]. These advances have also greatly promoted research on silkworm functional and evolutionary genomics.

With re-sequencing data of representative domesticated silkworm strains and wild silkworm samples from diverse geographic regions, the candidates of domestication genes were identified [12]. However, one puzzling observation in this study is that, although domesticated strains are clearly genetically differentiated from the wild ones, they still harbor ~83% of the variation observed in wild silkworm [12]. Later, using genetic evidence and coalescent simulation method, we showed that the domesticated silkworm lost 33-49% of nucleotide diversity relative to wild silkworm, which may attribute to a historical bottleneck during its domestication [13]. Furthermore, we investigated the phylogeny and evolutionary history of silkworm based on both mitochondrial loci and nuclear loci sequences [14]. However, some issues about inference

* Correspondence: yshen1964@126.com

†Equal contributors

¹State Key Laboratory of Silkworm Genome Biology, The Key Sericultural Laboratory of Agricultural Ministry, Southwest University, Chongqing 400715, China

Full list of author information is available at the end of the article

of the demographic history involving gene flow of *B. mori* remain unresolved. Probably, the lack of sophisticated statistical methods limits the study of demographic history of silkworm. Fortunately, the latest Approximate Bayesian Computation (ABC) method is suitable for assuming different demographic models in Bayesian framework to infer evolutionary history of model species [15-17].

A growing number of studies have demonstrated that gene flow between wild and domesticated animal populations during domestication occurred frequently [18]. Typical evidence was found not only in pigs [19], sheep and goats [20], as well as Bactrian camel [21] but also in domesticated crop plants Genus *Zea* [1] and Pearl Millet [22]. However, the demographic history of silkworm involving gene flow has not been investigated. Because gene flow likely occurred in different phases of the whole demographic history of silkworm, in this study we first referred the evolutionary models of the crop species mentioned above and hypothesized different evolutionary models to describe the demographic history of *B. mori*. Then, we evaluated the most possible demographic scenario using the coalescent simulation and ABC method [23,24] by combining the available nuclear DNA sequences and our newly sequenced data. Additionally, we also discussed the introgression pattern based on inferred evolutionary scenario with the goal of understanding the demographic history of silkworm.

Results

The characteristics of data used in this study

To avoid the effect of artificial selection and natural selection on the inference of the demographic history of silkworm as well as the detection of gene flow, we chose 17 loci that show neutral evolution in our previous studies [13,25]. Furthermore, Tajima's test [26] on each of 12 newly sequenced loci is not significant (Additional file 1), suggesting that these loci evolved in a neutral manner. Thus, all of 29 loci used in this study show a neutral evolution pattern.

Demographic history of silkworm

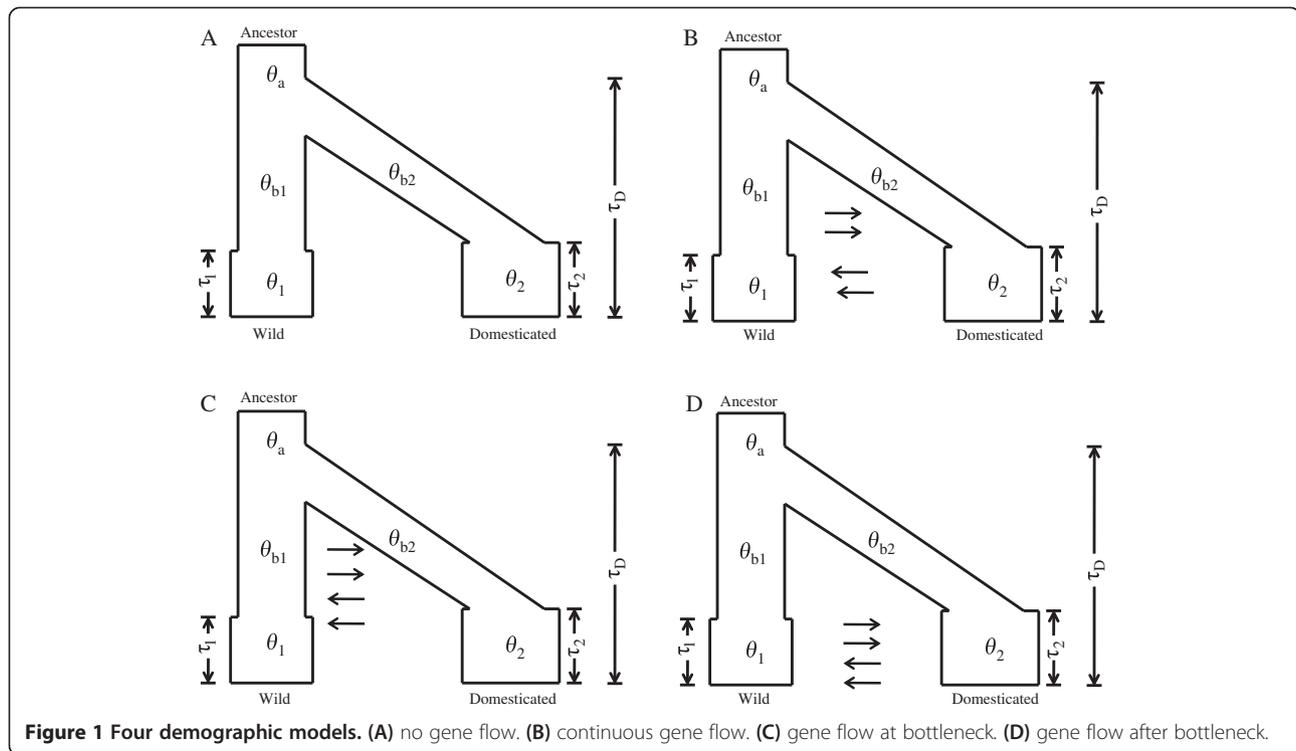
To investigate the occurrence of gene flow between the domesticated and wild silkworms, the demographic history of silkworm domestication was estimated by using coalescent simulation and ABC method. Four models ('no gene flow' model, 'continuous gene flow' model, 'gene flow at bottleneck' model and 'gene flow after bottleneck' model) were assumed (Figure 1). Eight demographic parameters (θ_1 , θ_2/θ_1 , θ_a/θ_1 , θ_{b2}/θ_1 , θ_{b1}/θ_1 , τ_D , τ_1 and τ_2) were estimated. The posterior density curve of each parameter for each model was also calculated (Additional files 2, 3, 4 and 5). These parameters can be converted into effective population size (N_e) and time in years using the formulation of $\theta = 4N_e\mu$ and in

units of $4N_1$ generations, assuming the mutation rate $\mu = 1.56 \times 10^{-8}$ /bp/generation [10]. The values of these parameters are calculated (Additional file 6). We found that four models have very different parameter values. For example, N_{b2} is equal to $\sim 10,793$ individuals (95% confidence interval (CI): 2,075 – 95,560) in the 'no gene flow' model, $\sim 9,310$ individuals (95% CI: 2,100 – 65,250) in the 'continuous gene flow' model, $\sim 7,669$ individuals (95% CI: 1,390 – 63,205) in the 'gene flow at bottleneck' model, and $\sim 8,279$ individuals (95% CI: 1,785 – 64,820) in the 'gene flow after bottleneck' model; T_D is equal to $\sim 6,334$ years (95% CI: 3,000 – 55,120) in the 'no gene flow' model, $\sim 7,424$ years (95% CI: 3,080 – 57,800) in the 'continuous gene flow' model, $\sim 7,460$ years (95% CI: 2,640 – 56,680) in the 'gene flow at bottleneck' model, and $\sim 5,906$ years (95% CI: 3,360 – 63,560) in the 'gene flow after bottleneck' model.

To determine which model is the most likely scenario for silkworm domestication, the value of Bayes factor (K) was calculated. The results showed that the maximum value of K was produced by 'gene flow at bottleneck' model (Table 1). According to this model, the scenario for silkworm domestication was as follows. The effective population sizes of the domesticated silkworm population at the modern and bottleneck stages are 73,196 (95% confidence interval (CI): 10,270 – 549,065) and 7,669 (95% CI: 1,390 – 63,205), respectively. Additionally, we also estimated the divergence time: the starting time of the domestication is $\sim 7,460$ years ago (95% CI: 2,640 – 56,680); the time of bottleneck ended in the domesticated silkworm is $\sim 3,984$ years ago (95% CI: 1,980 – 42,900) (Table 2). This model showed that gene flow occurred at the bottleneck stage during silkworm domestication. However, the direction of gene flow between the domesticated silkworm and wild silkworm is still unknown.

Detection of gene flow direction

To determine the direction of gene flow between the two populations, three models were assumed (Figure 2). Based on these models, simulations were performed by using the software of MIGRATE 3.2.16. The results showed that the model probabilities of Model A, Model B and Model C are 0.9999, 0.0001 and 0.0000, respectively (Table 3). This indicates that Model A with the bidirectional gene flow during bottleneck is the most likely scenario. Accordingly, we estimated the migration rates. The migration rate from wild to domesticated silkworms is 107.1 (95% CI: 58.3 ~ 153.8) (number of migrants per generation) while the migration rate from domesticated to wild silkworms is 318.6 (95% CI: 237.2 ~ 397.4). To further confirm bidirectional gene flow occurred during silkworm domestication, the IMA program was used to perform simulation analysis [27,28].



Likelihood-ratio test on bidirectional migration between ancestral sub-populations is significant ($P < 0.001$). The maximum likelihood estimates of the bidirectional migration rate from wild to domesticated silkworm population (m_1) and from domesticated to wild silkworm population (m_2) are 0.25 (95% CI: 0.0235 ~ 0.9735) and 0.91 (95% CI: 0.0352 ~ 1.4603) (per mutation), respectively (Additional file 7). This further suggests that the bidirectional migration might occur during silkworm domestication. Furthermore, the results of both approaches suggested that the migration rate from domesticated to wild silkworms is larger than that from wild to domesticated silkworms.

Table 1 Estimates of Bayes factor for hypothesized models

Statistic of demographic inference	Model	K
$\pi_{\text{silent-D}}$	A: No gene flow	1
	B: Continuous gene flow	80.5
	C: Gene flow at bottleneck	293.63
	D: Gene flow after bottleneck	8.8
$S_{\text{silent-D}}$	A: No gene flow	1
	B: Continuous gene flow	109
	C: Gene flow at bottleneck	185.26
	D: Gene flow after bottleneck	11.3

K: Bayes factor; $\pi_{\text{silent-D}}$: π_{silent} values for *B. mori*; $S_{\text{silent-D}}$: S_{silent} values for *B. mori*.

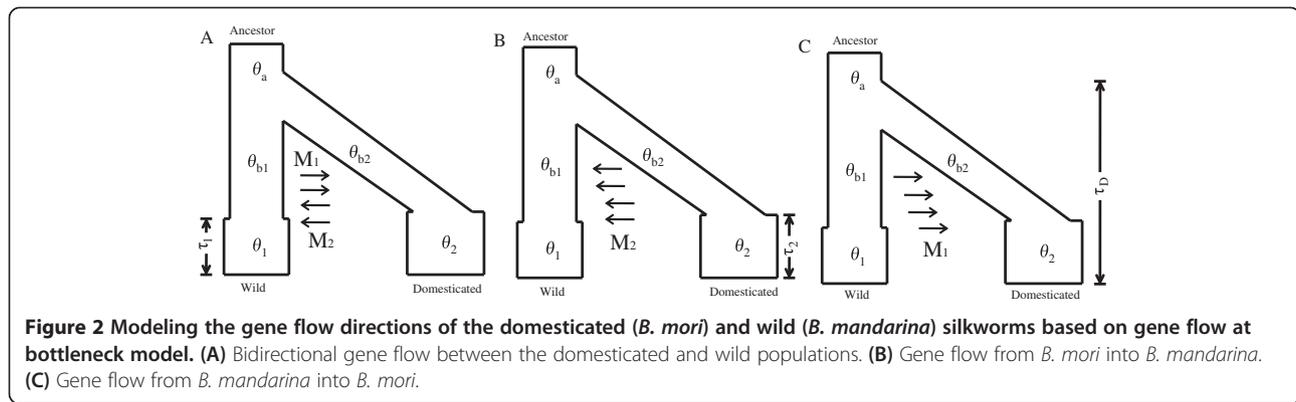
Discussion

Since the genome sequence of the domesticated silkworm was completed, many studies have used the genome-wide nuclear data and the classic sequencing data of individual nuclear loci to infer the demography and domestication of silkworm [12-14,25,29]. Although

Table 2 Silkworm domestication history parameters of preferred model

Parameters	Max	Lower 95%	Upper 95%
θ_a	0.98953	0.86365	1.16068
N_a	494,765	431,825	580,340
θ_2	0.146392	0.02054	1.09813
N_2	73,196	10,270	549,065
θ_{b1}	0.94133	0.11927	1.31742
N_{b1}	470,665	59,635	658,710
θ_{b2}	0.015339	0.00278	0.12641
N_{b2}	7,669	1,390	63,205
τ_D	0.00373	0.00132	0.02834
T_D	7,460	2,640	56,680
τ_1	0.0025	0.00107	0.02438
T_1	5,000	2,140	48,760
τ_2	0.001992	0.00099	0.02145
T_2	3,984	1,980	42,900

Parameters of the preferred model were estimated. θ and τ are converted to N and T (years) respectively. Conversion procedure is based on the equation $\theta = 4N\mu$, neutral mutation rate $\mu = 1.56 \times 10^{-8}$ per site per generation was used.



these studies provide a lot of useful information for understanding the origin and domestication of the domesticated silkworm, they appear to ignore detection of gene flow between the domesticated and wild populations.

Previous studies based on the nuclear genome data suggested that the domesticated silkworm lost only approximately 17% of nucleotide diversity measured by θ relative to wild silkworm [12]. However, our previous studies revealed that the domesticated silkworm lost 33-49% of nucleotide diversity measured by θ relative to wild silkworm and experienced a severe bottleneck during domestication [13,14,25]. As indicated previously, the main reason leading to this big difference is that studies based on Solexa sequencing data might have underestimated the level of nucleotide diversity in wild silkworm [13]. The nucleotide diversity in the domesticated silkworm was found to be at a very similar level measured as θ between the classic sequencing data and Solexa resequencing data, however, the nucleotide diversity in wild silkworm was greatly underestimated when the Solexa resequencing data were used [13]. The possible reason is that the genome sequence of wild silkworm, *B. mandarina*, has not been available to date, thus, the genome sequence of the domesticated silkworm was used as a reference for calling SNPs in both the domesticated and wild silkworms. Because resequencing reads are usually short (~45 bp), some mapping algorithms cannot map sequence reads with more than one or two differences from a reference genome sequence [30]. Thus, we decided to reconsider the demographic history of the domesticated silkworm using a combined data

set of nuclear loci by classic sequencing rather than the whole genome resequencing data by Solexa. It will be good time to better infer the demographic history of the domesticated silkworm using the whole genome resequencing data only when the genome sequence of wild silkworm, *B. mandarina*, is available in future.

Since previous studies have ignored the detection of gene flow between the domesticated and wild silkworms [12-14,25], this study intended to detect possible introgression pattern of the domesticated and wild populations. Currently, very few studies had investigated the evolutionary pattern of the domesticated silkworm, *B. mori*. Therefore, quite limited available model information could be used when we started to investigate the demographic history of the domesticated silkworm. Using previous studies on historical divergence and gene flow of genus *Zea* and Pearl Millet as references [1,22], therefore, we hypothesized four evolutionary scenarios by adding gene flow to reconsider the silkworm's domestication. We found that the 'gene flow at bottleneck' model is the most likely scenario. According to this model, the starting time of the silkworm domestication is ~7,460 years ago (95% CI: 2,640 – 56,680); the time of bottleneck ended in the domesticated silkworm is ~3,984 years ago (95% CI: 1,980 – 42,900). These results are consistent with our previous results based on the phylogenetic analyses in which the domestication of silkworm was estimated to occur about 4,100 years ago [14]. In addition, the estimate of the silkworm domestication time is also well consistent with the domestication time estimates of rice (~7,500 years) as well as maize (~7,500 years) [31,32]. These results are nearly

Table 3 Log marginal likelihood values, LBF and model probability of migration models

Model	Thermodynamic score	LBF	Relative model probability
D ← → W	-50024.87	0.00	0.9999
D → W	-50032.70	-7.83	0.0001
W → D	-50052.43	-19.82	0.0000

D: *B. mori*; W: *B. mandarina*; LBF: log Bayes factor.

consistent with the evidence from fossils [33]. In addition, the domestication time of most livestock can be traced back to 8,000-10,000 years ago [34]. Our divergence timing estimation is also within the timeframe. Therefore, our results are reasonable.

Coalescent simulation indicated that the gene flow might occur during the period of bottleneck. Three possible reasons may explain this. First, the overlapping geographic regions between the domesticated and wild silkworm populations might facilitate the gene flow [35]. Wild silkworm is not only widely distributed in the regions of Jiangsu and Zhejiang provinces, but also can be found in Anhui and Shandong provinces in China. Interestingly, these areas are also the main regions of sericulture [36]. Second, the lack of the absolute reproductive isolation is likely to be another important factor that leads to gene flow between the populations. In fact, the domesticated and wild silkworms have not formed the completely reproductive isolation up to now. They are able to interbreed with each other. This may provide some convenience for gene exchanging between these two species. Third, the environmental factors may also provide opportunities for gene flow between the populations during silkworm domestication. The initial rearing conditions of the domesticated silkworm may be very simple and open compared with present-day's rearing environments [36]. This may also increase the opportunities of contact and hybridization between the domesticated and wild silkworms. Besides these factors, economic expansion and urbanization might also affect gene flow. Rapid economic expansion and intensive urbanization may be common causes of the habitat fragmentation and may eventually reduce the gene flow on wildlife populations between different fragmented landscapes [37,38]. Thus economic expansion and urbanization could be the reasons of why gene flow occurred at the period of bottleneck rather than at the recent.

To determine the direction of gene flow during bottleneck, simulations were performed by using the software of MIGRATE 3.2.16. The results indicated that bidirectional gene flow during the bottleneck period between the populations is the most likely scenario (Table 3). In addition, the IM simulation results also supported the bidirectional gene migration between the domesticated and wild silkworm populations rather than the unidirectional gene flow model (Additional file 7). The gene flow direction from the domesticated silkworm to wild silkworm might be caused by the situation that the domesticated silkworms were discarded together with their excrement into natural environment [36]. The gene flow from wild silkworm to the domesticated silkworm might be because of that wild silkworm was introduced by breeders, as an important genetic source, to hybrid with

the domesticated silkworm to produce desirable strains [8,39].

In this study, we revealed that there was gene flow between the domesticated and wild silkworms. Our finding is consistent with the observations in maize [1], pig [19,40], chicken [41], cattle [42,43], cats [44], horses [45], and honeybee [46]. It appears that gene flow is ubiquitous between domesticated species and their wild relatives. Our results may add a dimension to highlight the important role of gene flow in domestication of crops and animals. Furthermore, as more and more population genomics data become available and new statistical methods are developed, we believe that long-standing questions about the role of gene flow in species formation will be resolved soon [47].

Conclusions

In this study, we hypothesized different models to describe the demographic history and the gene flow between *B. mori* and its wild relative, *B. mandarina*. The starting time of silkworm domestication was estimated to be approximate 7,500 years ago; the time of domestication termination was 3,984 years ago. Bi-directional gene flow might occur during silkworm domestication. Our findings add a dimension to highlight the important role of gene flow in domestication of crops and animals.

Methods

Data sources

In previous studies, we sequenced 17 loci in the domesticated silkworm (6 to 17 strains) and wild silkworm (7 to 15 samples) [13,25]. To expand our data set, we sequenced another 12 nuclear loci in this study, in total the current data set includes 29 loci DNA sequences (Additional file 8). The domesticated and wild silkworms used in this study are almost the same as those used in previous study [13,25]. All new sequences have been submitted to GenBank (Accession nos. KC137982-KC138225, KF703556-KF703614). The domesticated silkworm strains were randomly obtained from the Institute of Sericulture and Systems Biology at Southwest University, China, and represented the four main geographic strains (Chinese, Japanese, European and Tropical). Wild silkworm samples were collected from various geographical regions in China. No specific permissions were required for sampling wild silkworm. Sequences were aligned by BioEdit 7.0.4.1 [48]. For each locus, π_{silent} (π values for synonymous and noncoding sites) and θ_{silent} (θ values for synonymous and noncoding sites) were calculated using DNAsp 5.00 [49].

Demographic history analysis

The coalescent simulation was performed using ms software [50] and the four models ('no gene flow' model,

'continuous gene flow' model, 'gene flow at bottleneck' model and 'gene flow after bottleneck' model) were introduced (Figure 1). No gene flow model assumes no gene flow between the domesticated silkworm and wild silkworm populations. Continuous gene flow model assumes continuous gene flow between the populations. In the 'gene flow at bottleneck' model, gene flow occurred only in the bottleneck period. In the 'gene flow after bottleneck' model, gene flow occurred after the end of bottleneck period. For all four models, eight demographic parameters (θ_1 , θ_2/θ_1 , θ_a/θ_1 , θ_{b2}/θ_1 , θ_{b1}/θ_1 , τ_D , τ_1 and τ_2) were calculated (the ancestral population evolved with a population mutation rate θ_a and the ancestral population split into two bottlenecked populations of size θ_{b1} and θ_{b2} at time τ_D . The wild silkworm population θ_1 was assumed as reference population. The daughter populations remained small until they recovered from their bottlenecks to modern sizes of θ_1 and θ_2 at times τ_1 and τ_2 in the past). In our simulations, prior values and distribution of the parameters were listed in Additional file 9. The prior values were referenced the previous studies in model organisms [1,17]. For each model, 2,000,000 simulations were performed using the ms software [48]. To summarize the dataset, we made use of two statistics (the number of segregating sites at silent sites (S_{silent}) and the average number of pairwise differences at silent sites (π_{silent}) for each population) for demographic inference. The posterior probability distributions of eight parameters were calculated using rejection-regression procedure [15] and acceptance values of 0.001 were used for all analyses. Regression analysis was performed using the program ABCreg [51].

To determine which of the above four assumed models is the most likely scenario for silkworm domestication, Bayes factor was calculated using the following method [1]. Briefly, the values of the domesticated silkworm S and π were simulated using the program ms [50] with the parameter values drawn from the posterior distributions of the demographic models inferred by use of the 29 loci sequences. For each model, the acceptable simulation of each model was defined based on the Euclidean distance $\delta = 0.01$ ($|\text{simulated values} - \text{observed values}| < \delta$). Furthermore, the acceptance rate of each model was calculated. Finally, we assumed the isolation model as a reference model. The relative acceptance rates (the Bayes factor (K)) of each model were also calculated.

Detecting the direction of gene flow between populations

To determine the directional gene flow during bottleneck, MIGRATE 3.2.16 was used to perform analysis [52]. Three models based on inferred model were introduced. These models are shown in Figure 2. (1) Model A assumes two directional introgressions

(one from the domesticated silkworm into wild silkworm population and the other from wild silkworm into the domesticated silkworm population); (2) Model B assumes only unidirectional introgression (from domesticated silkworm into wild silkworm population); (3) Model C also assumes unidirectional introgression (from wild silkworm into the domesticated silkworm population). All parameters of the models are described in Additional file 10. For each model, four parallel static chains were performed with temperatures of 1.0, 1.5, 3.0 and 10^4 , with a number of 10,000 recorded steps and with 100 replicate runs. The log marginal likelihood values for the Bezier approximated score and the Harmonic mean were calculated by MIGRATE software. Furthermore, the probability of each model and log Bayes Factors (LBF) were calculated using the formulations as follows: (1) $\text{Prob}(\text{model}_i) = mL_{\text{model}_i} / \sum_j^n mL_{\text{model}_j}$; (2) $\text{LBF} = \ln(mL(\text{model}_1)) - \ln(mL(\text{model}_2))$. In addition, to further confirm the direction of gene flow during bottleneck, we also used the isolation with migration (IM) model as implemented in IMA software [27] to detect the possible gene flow directions between these two closely related populations. IM model also assumes the ancestral population splits into two descendant sub-populations whose sub-population can migrate with each other. IMA software is based on Markov Chain Monte Carlo (MCMC) to simulate gene genealogies. We performed Markov Chain Monte Carlo (MCMC) simulation with a million burn-in period followed by a million runs to generate genealogies. Our runs were conducted using the identical setting but with three different random number seeds. Furthermore, the estimates of joint-posterior densities and likelihood-ratio tests (LRT) were also performed to select the most likely model. Finally, based on IM model, the migration parameters from the wild to domesticated silkworm population (m_1) and the migration parameter from the domesticated silkworm to wild silkworm population (m_2) were also calculated.

Additional files

Additional file 1: Table S1. Summary statistics of nucleotide diversity and neutrality test results of 12 loci sequenced in this study.

Additional file 2: Figure S1. Posterior distributions of demographic parameters in no gene flow model.

Additional file 3: Figure S2. Posterior distributions of demographic parameters in continuous gene flow model.

Additional file 4: Figure S3. Posterior distributions of demographic parameters in gene flow at bottleneck model.

Additional file 5: Figure S4. Posterior distributions of demographic parameters in gene flow after bottleneck model.

Additional file 6: Table S2. Estimation values of parameters for hypothesized four demographic models.

Additional file 7: Figure S5. Maximum likelihood estimates of migration parameters using IMA analysis.

Additional file 8: Table S3. Summary of all loci information of *B. mori* and *B. mandarin*.

Additional file 9: Table S4. Priors used for Approximate Bayesian Computation.

Additional file 10: Table S5. Parameters of all hypothesized models.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YHS and ZZ designed the study. SY and MJH performed the analyses and drafted the manuscript. LFK and ZWL performed the experiments. YHS and ZZ supervised the study and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Dr. Fang-Yin Dai for help in collecting the domesticated silkworm samples and the members of Zhang's laboratory for their help with data analysis and helpful discussions. We also thank Peter Beerli for his help in data analysis and helpful discussion. This work was supported by the Hi-Tech Research and Development (863) Program of China (2013AA102507).

Author details

¹State Key Laboratory of Silkworm Genome Biology, The Key Sericultural Laboratory of Agricultural Ministry, Southwest University, Chongqing 400715, China. ²Laboratory of Evolutionary and Functional Genomics, School of Life Sciences, Chongqing University, Chongqing 400044, China.

Received: 14 April 2014 Accepted: 5 August 2014

Published: 14 August 2014

References

- Ross-Ibarra J, Tenaillon M, Gaut BS: Historical divergence and gene flow in the genus *Zea*. *Genetics* 2009, **181**:1399–1413.
- Larson G, Burger J: A population genetics view of animal domestication. *Trends Genet* 2013, **29**:197–205.
- Dempewolf H, Hodgins KA, Rummel SE, Ellstrand NC, Rieseberg LH: Reproductive isolation during domestication. *Plant Cell* 2012, **24**:2710–2717.
- Stevison LS, Kohn MH: Divergence population genetic analysis of hybridization between *rhesus* and *cynomolgus macaques*. *Mol Ecol* 2009, **18**:2457–2475.
- VonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkahoul A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, et al: Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 2010, **464**:898–902.
- Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F, Besnier F, Carlborg Ö, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L: Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 2010, **464**:587–591.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J, Li S, Larkin DM, Kim H, Frantz LAF, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Zhan B, et al: Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 2012, **491**:393–398.
- Goldsmith MR, Shimada T, Abe H: The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol* 2005, **50**:71–100.
- Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y: The genome sequence of silkworm, *Bombyx mori*. *DNA Res* 2004, **11**:27–35.
- Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, Zhao P, Zha X, Cheng T, Chai C, Pan G, Xu J, Liu C, Lin Y, Qian J, Hou Y, Wu Z, Li G, Pan M, Li C, Shen Y, Lan X, Yuan L, Li T, Xu H, Yang G, Wan Y, Zhu Y, Yu M, Shen W, et al: A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* 2004, **306**:1937–1940.
- International Silkworm Genome Consortium: The genome of a lepidopteran model insect, the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* 2008, **38**:1036–1045.
- Xia Q, Guo Y, Zhang Z, Li D, Xuan Z, Li Z, Dai F, Li Y, Cheng D, Li R, Cheng T, Jiang T, Becquet C, Xu X, Liu C, Zha X, Fan W, Lin Y, Shen Y, Jiang L, Jensen J, Hellmann I, Tang S, Zhao P, Xu H, Yu C, Zhang G, Li J, Cao J, Liu S, et al: Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). *Science* 2009, **326**:433–436.
- Guo Y, Shen Y, Sun W, Kishino H, Xiang Z, Zhang Z: Nucleotide diversity and selection signature in the domesticated silkworm, *Bombyx mori*, and wild silkworm, *Bombyx mandarina*. *J Insect Sci* 2011, **11**:1–16.
- Sun W, Yu H, Shen Y, Banno Y, Xiang Z, Zhang Z: Phylogeny and evolutionary history of the silkworm. *Sci China Life Sci* 2012, **55**:483–496.
- Beaumont MA, Zhang W, Balding DJ: Approximate Bayesian computation in population genetics. *Genetics* 2002, **162**:2025–2035.
- Francois O, Blum MG, Jakobsson M, Rosenber NA: Demographic history of european populations of *Arabidopsis thaliana*. *PLoS Genet* 2008, **4**:e1000075.
- Ross-Ibarra J, Wright SI, Foxe JP, Kawabe A, DeRose-Wilson L, Gos G, Charlesworth D, Gaut BS: Patterns of polymorphism and demographic history in natural populations of *Arabidopsis lyrata*. *PLoS One* 2008, **3**:e2411.
- Marshall FB, Dobney K, Denham T, Capriles JM: Evaluating the roles of directed breeding and gene flow in animal domestication. *Proc Natl Acad Sci U S A* 2014, **111**:6153–6158.
- Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresselt A, Vigne J-D, Edwards CJ, Schlumbaum A, Dinu A: Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci U S A* 2007, **104**:15276–15281.
- Naderi S, Rezaei H-R, Taberlet P, Zundel S, Rafat S-A, Naghash H-R, El-Barody MA, Ertugrul O, Pompanon F, Consortium E: Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS One* 2007, **2**:e1012.
- Ji R, Cui P, Ding F, Geng J, Gao H, Zhang H, Yu J, Hu S, Meng H: Monophyletic origin of domestic bactrian camel (*Camelus bactrianus*) and its evolutionary relationship with the extant wild camel (*Camelus bactrianus ferus*). *Anim Genet* 2009, **40**:377–382.
- Cloutault J, Thuillet AC, Buiron M, De Mita S, Couderc M, Haussmann BI, Mariac C, Vigouroux Y: Evolutionary history of pearl millet (*Pennisetum glaucum* [L.] R. Br.) and selection on flowering genes since its domestication. *Mol Biol Evol* 2012, **29**:1199–1212.
- Excoffier L, Estoup A, Cornuet JM: Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* 2005, **169**:1727–1738.
- Tanaka MM, Francis AR, Luciani F, Sisson S: Using approximate Bayesian computation to estimate tuberculosis transmission parameters from genotype data. *Genetics* 2006, **173**:1511–1520.
- Yu H, Shen Y, Yuan G, Hu Y, Xu H, Xiang Z, Zhang Z: Evidence of selection at melanin synthesis pathway loci during silkworm domestication. *Mol Biol Evol* 2011, **28**:1785–1799.
- Tajima F: Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989, **123**(3):585–595.
- Hey J, Nielsen R: Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci U S A* 2007, **104**:2785–2790.
- Hey J: On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol* 2005, **3**:e193.
- Li D, Guo Y, Shao H, Tellier LC, Wang J, Xiang Z, Xia Q: Genetic diversity, molecular phylogeny and selection evidence of the silkworm mitochondria implicated by complete resequencing of 41 genomes. *BMC Evol Biol* 2010, **10**:81.
- Li H, Ruan J, Durbin R: Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 2008, **18**:1851–1858.
- Zhu Q, Zheng X, Luo J, Gaut BS, Ge S: Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice. *Mol Biol Evol* 2007, **24**:875–888.
- Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS: The effects of artificial selection on the maize genome. *Science* 2005, **308**:1310–1314.

33. Yoshitake N: **Phylogenetic aspects on the origin of Japanese race of the silkworm, *Bombyx mori* L.** *J Seric Sci Jpn* 1968, **37**:83–87.
34. Bruford MW, Bradley DG, Luikart G: **DNA markers reveal the complexity of livestock domestication.** *Nat Rev Genet* 2003, **4**:900–910.
35. Bolnick DI, Fitzpatrick BM: **Sympatric speciation: models and empirical evidence.** *Annu Rev Ecol Syst* 2007, **38**:459–487.
36. Lv H: *Silkworm Sericulture in China (in Chinese)*. Shang Hai: Shanghai Science and Technology Press; 1991.
37. De Moura PA, Quek S-P, Cardoso MZ, Kronforst MR: **Comparative population genetics of mimetic *Heliconius* butterflies in an endangered habitat; Brazil's Atlantic Forest.** *BMC Genet* 2011, **12**:9.
38. Fleischer RC, Delaney KS, Riley SPD, Fisher RN: **A rapid, strong, and convergent genetic response to urban habitat fragmentation in four divergent and widespread vertebrates.** *PLoS One* 2010, **5**:e12767.
39. Nakamura T, Banno Y, Fujii H: **Genetics of the "wild silkworm translucent" mutant (ows) discovered in the progenies after the cross between the domesticated silkworm, *Bombyx mori*, and the wild mulberry silkworm, *Bombyx mandarina*.** *Int J Wild Silkmoth & Silk* 2001, **6**:7–10.
40. Ottoni C, Flink LG, Evin A, Geörg C, De Cupere B, Van Neer W, Bartosiewicz L, Linderholm A, Barnett R, Peters J: **Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics.** *Mol Biol Evol* 2013, **30**:824–832.
41. Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Strömstedt L, Wright D, Jungerius A, Vereijken A, Randi E: **Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken.** *PLoS Genet* 2008, **4**:e1000010.
42. Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JEO: **African pastoralism: genetic imprints of origins and migrations.** *Science* 2002, **296**:336–339.
43. Verkaar EL, Nijman IJ, Beeke M, Hanekamp E, Lenstra JA: **Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin?** *Mol Biol Evol* 2004, **21**:1165–1170.
44. Pierpaoli M, Biro Z, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E: **Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary.** *Mol Ecol* 2003, **12**:2585–2598.
45. Jordana J, Parés P, Sánchez A: **Analysis of genetic relationships in horse breeds.** *J Equine Vet Sci* 1995, **15**:320–328.
46. Harpur BA, Minaei S, Kent CF, Zayed A: **Management increases genetic diversity of honey bees via admixture.** *Mol Ecol* 2012, **21**:4414–4421.
47. Sousa V, Hey J: **Understanding the origin of species with genome-scale data: modelling gene flow.** *Nat Rev Genet* 2013, **14**:404–414.
48. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1999, **41**:95–98.
49. Librado P, Rozas J: **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**:1451–1452.
50. Hudson RR: **Generating samples under a Wright–Fisher neutral model of genetic variation.** *Bioinformatics* 2002, **18**:337–338.
51. Thornton K: **Automating approximate Bayesian computation by local linear regression.** *BMC Genet* 2009, **10**:35.
52. Beerli P: **Comparison of Bayesian and maximum-likelihood inference of population genetic parameters.** *Bioinformatics* 2006, **22**:341–345.

doi:10.1186/s12862-014-0185-0

Cite this article as: Yang et al.: Demographic history and gene flow during silkworm domestication. *BMC Evolutionary Biology* 2014 **14**:185.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

