MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 4094–4111 doi: 10.1111/mec.13300

Population genetic structure and approximate Bayesian computation analyses reveal the southern origin and northward dispersal of the oriental fruit moth *Grapholita molesta* (Lepidoptera: Tortricidae) in its native range

SHU-JUN WEI,* LI-JUN CAO,* YA-JUN GONG,* BAO-CAI SHI,* SU WANG,* FAN ZHANG,* XIAO-JUN GUO,* YUAN-MIN WANG* and XUE-XIN CHEN†

**Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China, †Institute of Insect Sciences, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China*

Abstract

The oriental fruit moth (OFM) *Grapholita molesta* is one of the most destructive orchard pests. Assumed to be native to China, the moth is now distributed throughout the world. However, the evolutionary history of this moth in its native range remains unknown. In this study, we explored the population genetic structure, dispersal routes and demographic history of the OFM in China and South Korea based on mitochondrial genes and microsatellite loci. The Mantel test indicated a significant correlation between genetic distance and geographical distance in the populations. Bayesian analysis of population genetic structure (BAPS) identified four nested clusters, while the GENELAND analysis inferred five genetic groups with spatial discontinuities. Based on the approximate Bayesian computation approach, we found that the OFM was originated from southern China near the Shilin area of Yunnan Province. The early divergence and dispersal of this moth was dated to the Penultimate glaciation of Pleistocene. Further dispersal from southern to northern region of China occurred before the last glacial maximum, while the expansion of population size in the derived populations in northern region of China occurred after the last glacial maximum. Our results indicated that the current distribution and structure of the OFM were complicatedly influenced by climatic and geological events and human activities of cultivation and wide dissemination of peach in ancient China. We provide an example on revealing the origin and dispersal history of an agricultural pest insect in its native range as well as the underlying factors.

Keywords: China, dispersal, microsatellite, mitochondrial gene, peach

Received 15 August 2013; revision received 22 June 2015; accepted 29 June 2015

Introduction

Improved knowledge regarding the history of individual species will improve our ability to predict how they will react to environmental dynamics in the future and can promote strategies in species management and conservation (Porretta *et al.* 2007; Lyons *et al.* 2012; Wei *et al.* 2013). Climate change and dispersal are two major factors influencing the contemporary distribution and genetic diversity of organisms (Hewitt 2004; Broquet & Petit 2009). In the climate oscillation of the Quaternary, species retreated into refugia during the glacial period and then expanded when conditions became suitable (Hewitt 2000). The repeated changes in the ranges

Correspondence: Shu-Jun Wei, Fax: +86 10-51503899; E-mail: shujun268@163.com and Xue-Xin Chen, Fax: +86 571-88982868; E-mail: xxchen@zju.edu.cn

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

initially shaped the distribution pattern of most living species (Hewitt 2004). In recent times, demographic events, such as stepping-stone expansion (Kerdelhué *et al.* 2009), human-related dispersal (Parker *et al.* 2001; Halpern *et al.* 2008) and active migration (Lyons *et al.* 2012; Wei *et al.* 2013) of the species, interactively confounded their distribution ranges (Hewitt 2000). In contrast, the present genetic constitution of the populations and species, carrying consequent signals of these past dynamics, can be used to interpret the evolutionary history of species (Hewitt 2004).

The oriental fruit moth (OFM) *Grapholita molesta* (Busck, 1916) (Lepidoptera: Tortricidae) is a serious pest that inflicts severe damage to orchards worldwide. The larvae damage the shoots and fruits of the stone and pome plants belonging to the family Rosaceae, mostly in the genera *Prunus* and *Pyrus* (Rothschild & Vickers 1991). This includes many of the economically important fruits, such as peach, apple and pear (Cory 1925; Myers *et al.* 2007; Pinero & Dorn 2009).

Assumed to be native to China, the OFM is now distributed throughout the stone-fruit-growing areas of Asia, the Middle East, Europe, Africa, South and North America, New Zealand and Australia (Rothschild & Vickers 1991). The OFM was first described under the name of *Laspeyresia molesta* in the USA in 1916 (Quaintance & Wood 1916), at which time no record of this species was found in Europe. Prior to this, the moth was mentioned as an unidentified pest of the peach in Japan (1901) and Australia (1910). The record of this species in other regions of the globe, such as South Africa, South America and New Zealand, came even later, confirming a recent worldwide introduction of this species (Table S1, Supporting information).

Due to its economic importance and recent introduction to other continents, the genetic structure and dispersal routes of the OFM population have been extensively examined (Timm *et al.* 2008; Torriani *et al.* 2010; Kirk *et al.* 2013; Zheng *et al.* 2013). Global study of the OFM indicated that the Australian and Japanese populations clustered with populations from China. The number of private alleles was significantly higher among Chinese populations than among populations from Japan, Europe and North America (Kirk *et al.* 2013). Although the OFM adults displayed a weak flight ability (Yetter & Steiner 1932), no pattern of isolation by distance was found within South Africa or Italy (Timm *et al.* 2008; Torriani *et al.* 2010), which is usually typical of recently colonized populations (Endersby *et al.* 2006; Yang *et al.* 2012). It was presumed that active and passive dispersal associated with accidental anthropogenic displacements helped expand their geographical range (Torriani *et al.* 2010). The effect of orchard management practices (Torriani *et al.* 2010) and selective host switch (Zheng *et al.* 2013) might have played important roles in the observed patterns of genetic variation.

However, the spatial expansion and temporal variation of the OFM in its presumed native area of China remain unclear. The OFM has been considered to be originated in northwest China, but questioned by the wider native range of its plant hosts in central Asia (Rothschild & Vickers 1991). Compared to the pattern of expansion and its influencing factors observed on other continents in which the OFM was introduced within the past 100 years (Quaintance & Wood 1916), the evolutionary history of this moth in China might be affected in a more complicated manner via both ancient and recent events. For example, Pleistocene climatic fluctuations on the distributions and the genetic divergence were detected in several widespread species in China (Meng *et al.* 2008; Zhang *et al.* 2008). It is also possible that the changes in environmental conditions could lead to expansion of multiple post-glacial colonization events in the agro-ecosystem (Kobayashi *et al.* 2011).

Given the limited knowledge on evolutionary history of the OFM in its native range, we examined the population genetic structure of the OFM in China and South Korea using both mitochondrial DNA sequences and microsatellite loci. The origin and dispersal history of this moth was deciphered using the approximate Bayesian computation (ABC) approach, which has proved to be a powerful method in dealing with complicated scenarios (Beaumont *et al.* 2002; Bertorelle *et al.* 2010). Our research demonstrated the southern origin and a northward dispersal of an important agricultural insect in its native range of China as well as the underlying factors that contributed to shape the current patterns of population genetic structure and diversity.

Materials and methods

Specimen collection

We collected 754 OFM larvae from peach orchards at 20 locations across China and one location in South Korea between 2010 and 2012 (Table 1, Fig. 1). Ten individual *Grapholita dimorpha* were used for gene sequencing as out-groups of molecular dating. All specimens were preserved in absolute ethanol and stored at −80 °C prior to DNA extraction.

DNA extraction, mitochondrial gene sequencing and microsatellite genotyping

Total genomic DNA was extracted from one abdominal segment of individual larvae using the DNeasy Blood and Tissue Kit (Qiagen, Germany).

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Sample collection information of the *Grapholita molesta* used in this study

Group	Population	Collection location	Longitude (E)	Latitude (N)	Collection date (Month/Year)	Number of individuals
SL	*SL*	Yunnan Province, Shilin	103°19′48.00″	24°51′48.96″	May/11	17
	KD	South Korea, Daegu	128°36′05.20″	35°52′17.17″	Aug/12	10
CD	*CD*	Sichuan Province, Chengdu	104°03′53.48″	30°39′30.96″	Sep/10	53
ND	*ND*	Fujian Province, Ningde	119°32′52.56″	26°39′56.22″	Jul/11	30
CE	NB	Zhejiang Province, Ningbo	121°32′38.36″	29°52′06.01″	Sep/10	13
	WH	Hubei Province, Wuhan	114°18′19.29″	30°35′35.11″	Jul/11	28
	NJ	Jiangsu Province, Nanjing	118°47′48.76″	32°03′36.92″	Jul/10	30
	TS	Gansu Province, Tianshui	105°43′29.81″	34°34′51.10″	Apr/10	30
	YL	Shanxi Province, Yangling	108°05′05.04″	34°16′19.62″	Aug/10	39
	ZZ	Henan Province, Zhengzhou	113°37′31.32″	34°44′47.76″	Aug/10	30
	YA	Shanxi Province, Yanan	109°29′22.69″	36°35′07.53″	Aug/10	31
NO	TG	Shanxi Province, Taigu	112°33′04.89″	37°25′16.71″	Sep/10	36
	QD	Shandong Province, Qingdao	120°22′57.01″	36°04′01.99″	Jun/10	59
	SJ	Hebei Province, Shijiazhuang	114°30′53.50″	38°02′32.31″	May/10	47
	BD	Hebei Province, Baoding	115°27′53.30″	38°52′26.01″	Jul/10	36
	DX	Beijing City, Daxing	116°20′29.02″	39°43′36.94″	Aug/10	48
	MT	Beijing City, Mentougou	116°06′06.13″	39°56′25.23″	Jul/10	48
	PG	Beijing City, Pinggu	117°07′17′.03′	40°08′26.16″	Jun/10	48
	YQ	Beijing City, Yanqing	115°58′29.92″	40°27′23.97″	Jul/10	43
	DL	Liaoning Province, Dalian	121°36′52.86″	38°54′50.41″	Jul/10	48
	SY	Liaoning Province, Shenyang	123°25′53.29″	41°48′20.59″	Sep/10	30

Groups are divided based on the GENELAND analysis using mitochondrial genes. CE indicates a central group; NO indicates a northern group. The italic populations are used in DIYABC analyses.

For the mitochondrial markers, DNA sequences with a total length of 4223 bp from six of the 13 mitochondrial protein-coding genes (*cox1*, *cox2*, *atp8*, *atp6*, *nad5* and *nad1*) and one of the two mitochondrial ribosomal RNA genes (large ribosomal RNA gene, *rrnl*) were used in this study. Most of the seven gene sequences have been previously shown to be powerful markers in population genetic studies (Meng *et al.* 2008; Ma *et al.* 2012). The amplification and sequencing primers were designed using PRIMER PREMIER software version 5 according to the complete OFM mitochondrial genome (Gong *et al.* 2012) (Table S2, Supporting information). Polymerase chain reaction (PCR) was conducted using the Mastercycler pro system (Eppendorf, Germany) under the following conditions: an initial denaturation for 3 min at 94 °C, followed by 35 cycles of 10 s at 94 °C, 15 s at 48 °C and 2 min for the region of *cox1-atp6* and *nad5* and 1 min for the regions of *nad1* and *rrnL* at 68 °C, and a subsequent final extension for 10 min at 68 °C. The PCR components were added as recommended by the manufacturer of Takara LA *Taq* (Takara Biomedical, Japan). Amplified products were purified and sequenced directly from both strands using the ABI 3730xl DNA Analyzer (Applied Biosystems, USA).

For the nuclear markers, eight microsatellite loci previously isolated by Torriani *et al.* (2010) were genotyped in the study (Table S3, Supporting information). All loci were fluorescently labelled and amplified as previously described by Schuelke (2000). The size of the amplified PCR products was determined using the ABI 3730xl DNA Analyzer with GeneScan 500 LIZ size standard. Allele designation was obtained using the software GENEMAPPER version 4.0 (Applied Biosystems, USA).

Data preparation and genetic diversity analyses

The sequencing results of mitochondrial genes determined from both strands were assembled in SEQMAN within the LASERGENE suite version 7.1.2 (DNASTAR, Inc., USA). Each gene was aligned independently using CLUSTALW (Thompson *et al.* 1994) implemented in MEGA version 5 (Tamura *et al.* 2011) using the default set of parameters. Alignment of the nucleotide sequences of the protein-coding genes was guided by amino acid alignment. The standard diversity indices and background selection using Tajima's D (Tajima 1989) were calculated using ARLEQUIN suite version 3.5 (Excoffier & Lischer 2010).

The microsatellite loci determined from GENEMAPPER were first checked for stuttering and large allele dropout using MICRO-CHECKER version 2.2.3 (van Oosterhout *et al.* 2004). Null allele frequencies at each locus were calculated using the R package GENELAND version 4.0.4 (Guillot *et al.* 2005). The number of alleles and observed

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

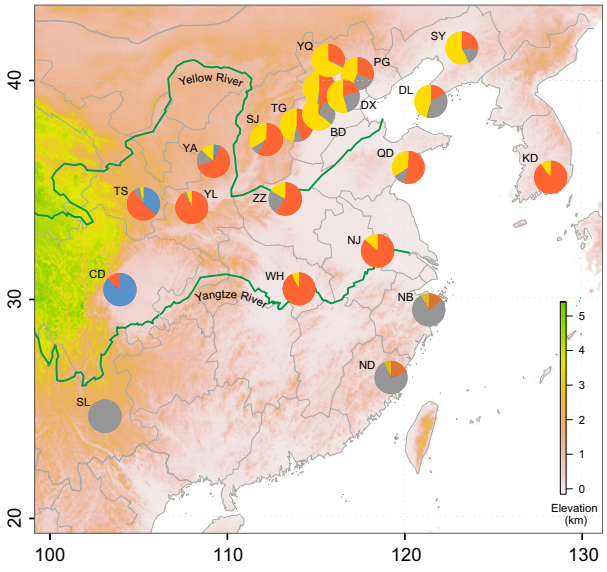


Fig. 1 Collection locations of the *Grapholita molesta* and spatial clustering of the individuals based on the combined mitochondrial genes estimated from BAPS. Detailed information on the locations and codes of the populations is presented in Table 1. The areas of different colours in each population correspond to the proportion of each cluster in the population. The blue cluster was mainly distributed in the CD population, reductively distributed in the TS and YA populations. The grey cluster was widely distributed in most of the sampled populations and decreased in proportion from southern to northern populations. The red cluster was mainly distributed in central China, and the yellow one was mainly distributed in northern China.

and expected heterozygosities were calculated using MICROSATELLITE ANALYSER (MSA) version 4.0.5 (Dieringer & Schlötterer 2003). Tests for linkage disequilibrium and deviation from Hardy–Weinberg equilibrium at each locus and for each population, and estimation of inbreeding coefficients (F_{IS}) for each population, were performed using GENEPOP version 4.0.11 (Raymond & Rousset 1995).

Haplotype relationship analysis

To reveal the relationships among haplotypes, split networks were constructed using the software SPLITSTREE version 4.13.1 (Huson & Bryant 2006) from both single and the combined mitochondrial genes. This software provides a framework for analysis using both trees and networks. The network methods in the software also provide a valuable tool for phylogenetic inference, even when reticulation events do not play an important role. To detect significant signature of reticulation, we used the neighbour-net method for construction of networks under a distance model of K2P. The statistically significant split with >95% confidence was identified after 1000 bootstraps.

Divergent time estimation

The divergence times for the two haplotype lineages were estimated using the software BEAST version 1.8.1 (Drummond & Rambaut 2007) based on the combined mitochondrial genes. The substitution model and the best partitioning schemes were determined simultaneously using the software PARTITIONFINDER version 1.1.1 (Lanfear *et al.* 2012) under a BIC and linked branch lengths. The data were partitioned into two partitions, which are composed of the protein-coding genes and the ribosomal RNA gene, respectively. The substitution model HKY + I + G for the protein-coding genes and HKY + G for the ribosomal RNA gene was applied. Coalescent tree priors were set to the constant size model. A relaxed clock with uncorrelated lognormal distribution was used to allow rate variation among branches, selected by the path sampling and stepping-stone sampling approaches (Baele *et al.* 2012, 2013). We applied the insect molecular clock for the combined *cox1* and *rrnl* genes (substitution rate = 1.345% per million years) (Papadopoulou *et al.* 2010), leaving the substitution rates for the other genes to be estimated. Two runs were executed for 200 million steps, sampling every 20 000 steps and discarding the initial 25% as burn-in. Convergence of the chains was checked using the program TRACER version 1.6 (http://tree.bio.ed.ac.uk/software/tracer) to ensure that effective sample sizes were above 200. Samples from the two runs, which yielded similar results, were combined for the estimation of the MRCA (most recent common ancestor) of the two lineages.

Population genetic structure analyses

To gain insight into the population genetic structure of the OFM, we used multiple approaches. Initially, pairwise F_{ST} between each pair of the sampled populations was calculated in the ARLEQUIN suite version 3.5 (Excoffier & Lischer 2010) for both mitochondrial genes and microsatellite loci.

Next, we performed Mantel tests of the genetic distance [$F_{ST}/(1\text{-}F_{ST})$] vs. the geographical distance (ln km) across populations for both mitochondrial genes and microsatellite loci using the software ZT version 1.1 (Bonnet & Van de Peer 2002) to estimate the level of isolation by distance. We used the F_{ST} values of microsatellite loci calculated by FREENA software, which uses an excluding null alleles (ENA) method to avoid the effect of null allele on the estimates of genetic differentiation (Chapuis & Estoup 2007).

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The spatially explicit BAPS model for clustering of individuals implemented in BAPS version 6.0 (Cheng *et al.* 2013) was used for mitochondrial genes. This model combines sample locations with the likelihood of the genetic data and is particularly efficient with large data sets. As the *K* values (maximum number of genetically diverged groups) affect the initial assignment of simulation and thus the possibility of finding only a local mode is reduced when the simulation is started many times from different initial assignments, we performed 20 runs ($K = 20$) to ensure convergence and consistency of the results.

Finally, the Bayesian clustering method implemented in the R package GENELAND version 4.0.4 (Guillot *et al.* 2005) was used to detect its geographical discontinuities for both mitochondrial genes and microsatellite loci. This method distributes the individual genotypes into K clusters by minimizing Hardy–Weinberg disequilibrium and gametic phase disequilibrium within the groups, taking into account the spatial coordinates. As the correlated model seems to be more prone to algorithm instabilities and more sensitive to departure from model assumptions, such as the presence of isolation by distance, we used the uncorrelated model of allele frequencies for both mitochondrial genes and microsatellite loci. The null allele model was allowed for microsatellite loci (Guillot *et al.* 2008). Four independent runs were performed with 10 million Markov chain Monte Carlo iterations, of which every 1000th one was saved. Initially, the tested number of genetic clusters (K) was set to vary between 1 and 21. To refine the geographical map of the genetic clusters, the model was run four more times, treating the number of clusters as known, based on the number of clusters, K, inferred from the previous runs. The posterior probability of population membership was computed using a burn-in period length of 300 iterations.

ABC-based Inferences about the origin and dispersal of the OFM: overall methodology

The ABC method, as implemented in the DIYABC version 2.0 program (Cornuet *et al.* 2014), was used to compare different competing scenarios regarding the ancestral populations of the OFM as well as the dispersal routes. This approach simulates the data sets for a number of predefined scenarios and compares the summary statistics of these with the summary statistics of the observed data, making it feasible to test complex population genetic models. The approach also allows for an admixture of two populations, which is suitable for analysing our data with nested population structures as inferred from BAPS analysis.

In the ABC analyses, mitochondrial genes and microsatellite loci were separately analysed in two sets of analyses. Details for the analyses based on the microsatellite loci are described in supporting information. We present the set of analyses based on the mitochondrial genes. Prior distributions were uniform and set as in Table S4 (Supporting information). The HKY model for the nucleotide substitution was used with the gamma value and the proportion of invariable sites estimated using the maximum-likelihood method in PHYML version 3.0 (Guindon *et al.* 2010). The 'one sample summary statistics' used were the number of haplotypes and the number of segregation sites, while the 'two-sample summary statistics' used was F_{ST}. Optimal number of simulated data sets was created for each analysis as recommended by the software.

The DIYABC program allows for estimating the posterior probability of the different scenarios compared. We compared the posterior probability of the competing scenarios using a polychotomic weighted logistic regression (Fagundes *et al.* 2007) with linear discriminant analysis (Estoup *et al.* 2012). Only 1% of the simulated data sets closest to the observed data were retained for these estimations (Beaumont *et al.* 2002). Confidence in the scenario choice was evaluated by computing type I error (risk to exclude the focal scenario when it is the true one) and type II error (risk to select the focal scenario when it is false) in the selection of scenarios. Posterior model checking was performed on the selected scenario of every analysis using a local linear regression on 1% of the simulated data sets closest to our real data (Beaumont *et al.* 2002). All summary statistics were used for model checking.

Scenarios of the ABC analyses

For the scenario test, the sampled individuals were treated as five geographical and genetic groups, as detected by the package GENELAND, named CD, SL, ND, CE (Central group) and NO (Northern group) (Table 1), to reduce the number and complexity of evolutionary scenarios to be compared, as in previous studies (Lombaert *et al.* 2014). To balance the number of analysed individuals among the five groups, the individuals from the sampled populations of TS, YL and WH, representing the CE group, and those from DL and SY, representing the NO group, were selected for analysis, rather than including all individuals identified in the two GENELAND groups. All of the selected populations are those with the highest posterior probabilities of population membership inferred by GENELAND. Because as many as 120 scenarios can be tested with five populations, we narrowed this number by defining nested subsets of competing scenarios that were analysed sequentially, based

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

on genetic diversity, haplotype relationships and population structure.

In step 1, the relationships among the three populations of CD, SL and ND with high genetic diversity, which were adjacently located in the southern region of China, were investigated. A total of six possible scenarios among the three populations were tested, regarding the variation in population size and the split and admixture events (Step 1-1 in Fig. 2). In this analysis, the size of ancestral population was varied in all scenarios. Subsequently, two competing scenarios based on the scenario that ND is from the admixture of CD and SL were analysed, regarding the source of the southern populations (Step 1-2 in Fig. 2). When the variation in the population size was allowed for the ancestral population in this analysis, the two scenarios could not be discriminated with equal probability. Thus, the population size was set to be constant in the analysis.

In step 2, we supposed three major hypotheses on the origin of the OFM, that is southern, central and northern origin. In the scenario definition, the model of stepping-stone dispersal was assumed, based on the weak flight ability of the OFM (Yetter & Steiner 1932) and the presence of isolation by distance in the sampled populations. The ND population was set to be an admixture of SL and CD, while the CD to be a split of SL, as revealed in step 1. According to the results of demographic analysis in the following Bayesian skyline plot analysis, we set the population size of the NO to be varied and that of the others to be constant. Totally, six, two and one possible scenarios for the hypotheses of southern, central and northern origin of the OFM could be defined. In this step, we compared the competing scenarios of southern and central origin in two analyses (step 2-1 and step 2-2 in Fig. 2).

In step 3, the best-fit scenarios of southern and northern origins inferred from step 2 as well as the single scenario of northern origin of the OFM were compared (step 3 in Fig. 2).

Demographic history analysis

Demographic history was analysed for each of the five groups identified via the package GENELAND. To avoid the influence of multiple differentiated lineages within a population, the two haplotype lineages in the CD and CE groups identified by SPLITSTREE were analysed separately as in Zlojutro *et al.* (2006). Historical demographic trends were investigated through a mismatch distribution analysis (Slatkin & Hudson 1991; Schneider & Excoffier 1999) using ARLEQUIN suite version 3.5. We also investigated the population history by estimating the changes in the effective population size over time using a Bayesian skyline plot (Drummond *et al.* 2005) implemented in the software BEAST version 1.8.0 (Drummond & Rambaut 2007). This approach incorporates uncertainty in the genealogy using MCMC integration under a coalescent model. The piecewise-linear skyline model was selected for the Bayesian skyline coalescent tree priors. Chains were run for 200 million generations with a sampling of every 20 000 generations. The convergence and output of BEAST were checked and analysed using TRACER version 1.6.

Results

Genetic variation and neutrality test

For the combined mitochondrial genes, we observed 197 haplotypes from the 21 OFM populations (GenBank Accession nos: KF163966-KF164208, KT004678-KT004683). High haplotype diversity was observed both among genes (ranging from 0.208 to 0.827) (Table S5, Supporting information) and among the sampled populations (ranging from 0.721 to 0.966) (Table 2). The populations from the southern and eastern regions had more private haplotypes than those from the other regions (e.g. 100.0%, 78.9%, 80.6% and 73.3% haplotypes in SL, ND, CD and TS are private in population) (Fig. S1, Supporting information). Tajima's D was not significantly different from 0 in all of the populations analysed (Table 2), which indicated that mitochondrial genes used in the study were under neutral equilibrium.

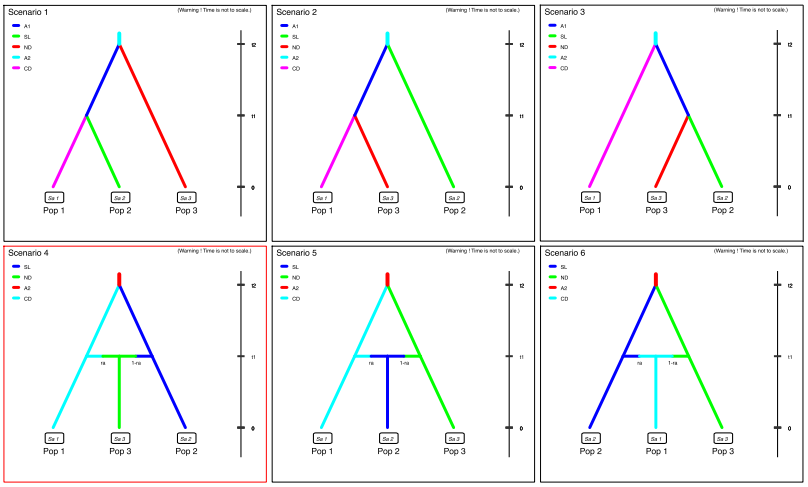
For the microsatellites, we included 14 populations for analyses, taking into account the frequency of null alleles and balance of sampling from the northern China (Table S3, Supporting information). There were 141 alleles in eight microsatellite loci. The number of alleles in an individual population ranged from 37 to 83. The observed heterozygosity in an individual population ranged from 0.2797 to 0.3706, and the expected heterozygosity ranged from 0.5566 to 0.7440 (Table S6, Supporting information). Significant linkage disequilib-

Fig. 2 Scenarios for the DIYABC analyses to infer the origin and dispersal of the *Grapholita molesta* based on mitochondrial genes. Step 1-1: six scenarios for relationships of three southern populations of CD, SL and NO, regarding the variation in population size and the split and admixture events; step 1-2: two competing scenarios based on the formation of ND from the admixture of CD and SL; step 2-1: six scenarios of southern origin; step 2-2: two scenarios of the northern origin; step 3: the best-fit scenarios of southern and northern origins inferred from step 2 and the single scenario of northern origin. The best scenario in the corresponding step was indicated by red square.

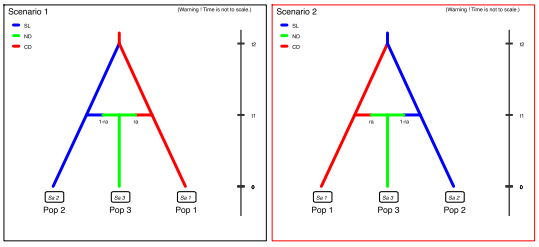
© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

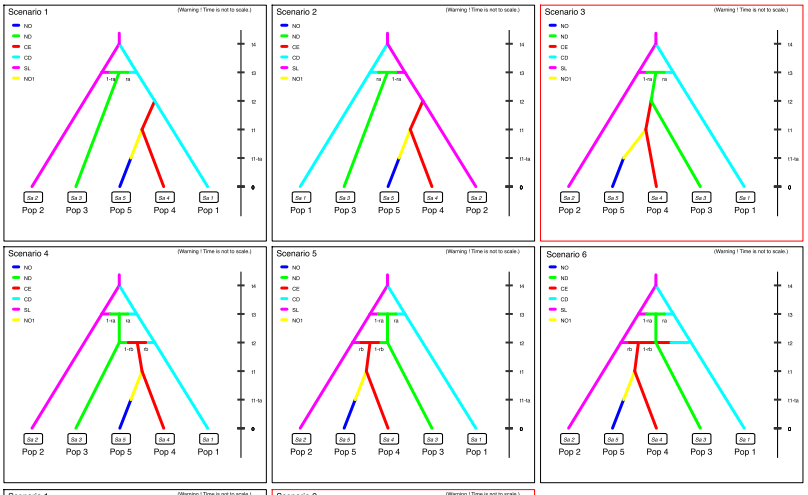
Step 1-1: six scenarios for relationships of three southern populations of CD, SL and NO, regarding the variation of population size and the split and admixture events.



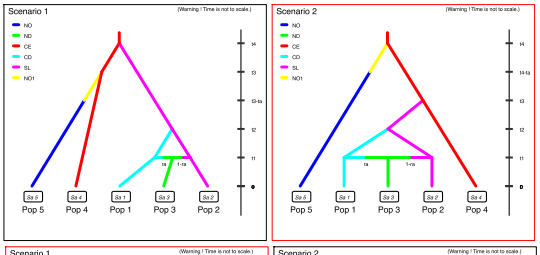
Step 1-2: two competitive scenarios based on the scenario that ND is from the admixture of CD and SL.



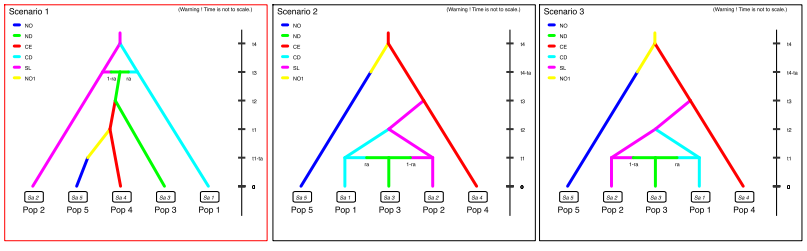
Step 2-1: six scenarios of southern origin.



Step 2-2: two scenarios of the northern origin.



Step 3: the best-fit scenarios of southern and northern origin inferred from step 2 as well as the single scenario of northern origin.



© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2 Parameters of the genetic diversity and the demographic analysis of 21 populations of *Grapholita molesta* based on the combined mitochondrial genes

Population	*S*	*h*	π	*P*	*D* (*p*)	τ (95% CI)	P_{SSD}
SL	12	0.721	0.00112	4.706	1.22258 (0.89900)	9.80664 (0.00000, 14.177773)	0.13935509
ND	39	0.966	0.00154	6.471	−1.26022 (0.09500)	8.32812 (4.56641, 11.24414)	0.01186369
NB	26	0.936	0.00168	7.077	−0.52775 (0.33500)	9.60938 (4.75977, 12.48438)	0.09139810
WH	17	0.841	0.00064	2.704	−1.37471 (0.06900)	0.77344 (0.28906, 1.13281)	0.10329815
NJ	18	0.855	0.00092	3.890	−0.49570 (0.34900)	0.21094 (0.00000, 5.78125)	0.02791498
CD	67	0.957	0.00307	12.938	−0.43010 (0.39700)	3.34570 (0.35938, 20.33789)	0.00842072
TS	56	0.924	0.00406	17.115	0.78976 (0.84300)	31.25977 (0.81250, 38.64648)	0.02941639
YL	27	0.807	0.00084	3.530	−1.46110 (0.05300)	0.05078 (0.00000, 3.56445)	0.01749447
ZZ	26	0.876	0.00140	5.910	−0.29513 (0.46300)	10.75977 (4.13086, 14.23242)	0.04375913
YA	50	0.925	0.00188	7.927	−1.37033 (0.07200)	9.54492 (3.66211, 14.28711)	0.02317419
TG	34	0.897	0.00139	5.849	−1.01308 (0.16300)	8.91797 (3.66992, 11.92969)	0.01674987
QD	30	0.802	0.00124	5.216	−0.62482 (0.2980)	8.49023 (3.15039, 12.03711)	0.03890435
SJ	42	0.931	0.00141	5.954	−1.28787 (0.08300)	8.93750 (4.43750, 12.55273)	0.01162220
BD	37	0.778	0.00141	5.944	−1.18678 (0.10100)	11.55273 (0.57422, 16.22461)	0.03646585
DX	37	0.858	0.00134	5.645	−1.09904 (0.11100)	8.90625 (3.89258, 13.16992)	0.01831298
MT	20	0.822	0.00130	5.478	0.69176 (0.80300)	9.13672 (1.85742, 13.30664)	0.03598778
PG	47	0.949	0.00161	6.737	−1.29261 (0.06900)	9.31641 (5.07422, 12.28320)	0.01144836
YQ	26	0.873	0.00121	5.087	−0.42693 (0.38300)	10.61133 (1.50000, 14.27344)	0.03225109
DL	27	0.873	0.00157	6.604	0.17725 (0.65300)	10.70703 (4.94336, 13.77930)	0.04503887
SY	31	0.782	0.00132	5.549	−1.10628 (0.12900)	9.59180 (1.65430, 13.16602)	0.05675067
KD	14	0.778	0.00101	4.244	−0.65186 (0.26100)	7.51758 (1.33008, 10.76758)	0.08216871

Number of polymorphic sites (*S*), haplotype diversity (*h*), nucleotide diversity (π), average number of pairwise differences (*P*) in the analysed regions of the *Grapholita molesta*; neutrality tests: Tajima's D (*D*) and expansion (coalescence) time under the sudden expansion assumption in mutation generations (τ) with 95% confidence interval (CI). Significance values (*P*) of the parameters were evaluated with 1000 simulations; P_{SSD}: *P* value for sum of squared deviations (SSD).

rium was present in 22 of 210 pairs of loci within populations and none of 28 locus pairs across all populations, but unlikely due to physical linkage (Torriani *et al*. 2010). The inbreeding coefficients are very high, ranged from 0.397 to 0.627 (Table S6, Supporting information), which might be related to sampling of each population from a single orchard and the weak flight ability of this moth (Yetter & Steiner 1932), or even to the effect of orchard management practices (Torriani *et al*. 2010). Hardy–Weinberg tests revealed seven of the eight loci across all sampling sites and all of the 14 populations across multilocus exhibited significant deviation, which is similar to the result of other studies on this moth (Torriani *et al*. 2010; Zheng *et al*. 2013).

Haplotype relationships and divergence times

The SPLITSTREE result revealed two haplotype lineages (Fig. 3): a minor one, which contained most haplotypes of CD, partial of TS and a few of the YA population, and a major one, which was composed of haplotypes from all populations. No significant signature of reticulation (with >95% confidence) was detected after 1000 bootstraps.

In addition, we estimated the divergent time of the two lineages by molecular clock analysis of mitochondrial genes. The divergence time of the two lineages was dated to 254 thousand years ago (Ka), with a 95% highest posterior density (HPD) of 159–371 Ka. The time to the MRCA was 110 Ka with 95% HPD of 63–160 Ka for the minor and 104 Ka with 95% HPD of 61–157 Ka for the major lineage.

Population genetic structure

The pairwise F_{ST} difference indicated significant differentiation in 84 of the 91 population pairs based on microsatellite loci (Table S7, Supporting information) and in 200 of the 210 based on mitochondrial genes (Table 3). For the microsatellite loci, no significant differences were observed between pairwise F_{ST} and ENA-corrected pairwise F_{ST} for null alleles (*t*-test, $P = 0.0678$), suggesting that null alleles did not affect this analysis.

The Mantel test results produced an *r* value of 0.269 for microsatellite loci ($P = 0.0099$) and 0.350 for combined mitochondrial genes ($P = 0.0013$) (Fig. 4, Fig. S2, Supporting information), which indicated that a significant correlation was found between the genetic distance and geographical distance in the OFM populations.

BAPS analysis of the mitochondrial genes identified four clusters in the 21 populations, exhibiting nested population structures (Fig. 1). The first cluster was mainly

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

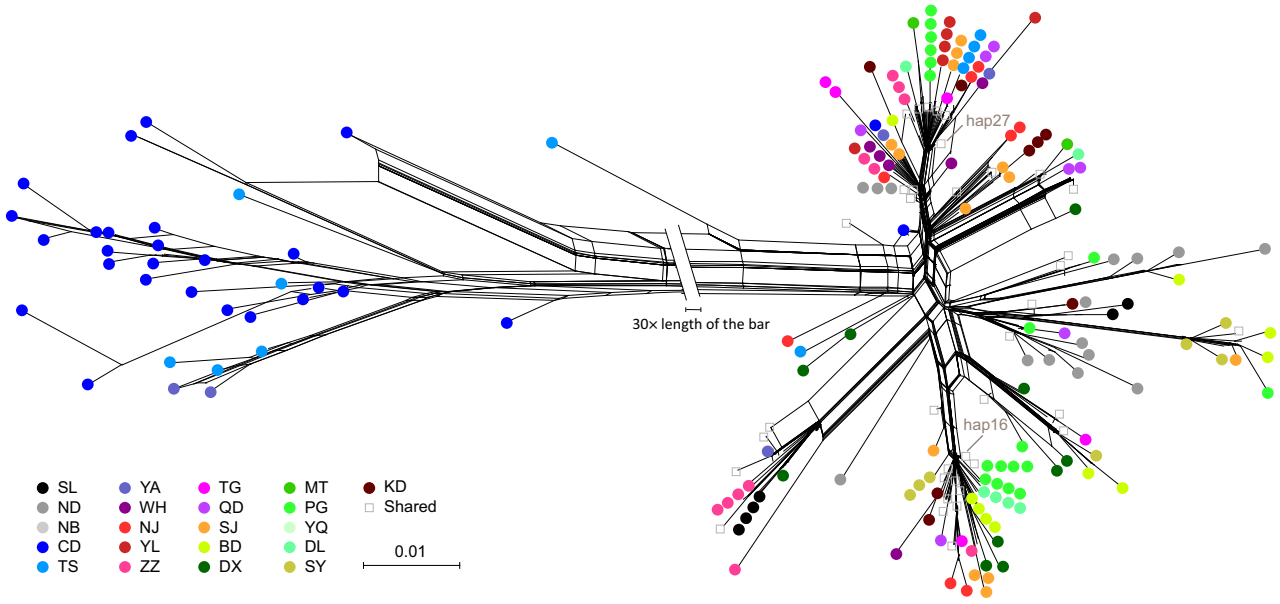


Fig. 3 Splits network for the *Grapholita molesta* haplotypes constructed by the method neighbour-net, based on combined mitochondrial genes. Circles of the same colour indicate haplotypes from the same population. 'Shared' indicates the haplotype shared by different populations. The Hap16 and Hap27 were the most commonly shared haplotypes.

distributed in the CD population, reductively distributed in the TS and YA populations, corresponding to the minor lineage in the SPLITSTREE tree (in blue in Fig. 1). The second cluster was widely distributed in most of the sampled populations and decreased in proportion from southern to northern populations (in grey in Fig. 1). The SL population was fully occupied by the second cluster. The third cluster was mainly distributed in central China (in red in Fig. 1), and the last one was mainly distributed in northern China (in yellow in Fig. 1).

Assuming uncorrelated allelic frequencies between sites, GENELAND inferred five distinct genetic groups based on mitochondrial genes, with spatial discontinuities reflecting geographical locations: a northern group, which encompassed populations located in the northern region of China (A1 in Fig. 5); a central group, which encompassed populations located in the central region of China (A2 in Fig. 5); and three small groups, which encompassed the populations of CD, SL + KD and ND (A3, A4 and A5 in Fig. 5). The microsatellite loci generated four groups (B1 to B4 in Fig. 5) corresponding to those inferred from the mitochondrial genes (A1 to A4 in Fig. 5). The lacking of ND group might be related to the lacking of microsatellite data in the sampling sites.

Origin and dispersal routes

Results for the scenario selection, evaluation of confidence and model checking of DIYABC analyses based on mitochondrial genes are summarized in Table S8, (Supporting information). Our evaluation of confidence in scenario choice revealed low type I and II errors in preliminary analyses (type I/II error of steps 1-1, 1-2, 2-1 and 2-2 is 0.22/0.16, 0.43/0.26, 0.26/0.30 and 0.10/0.10, respectively), and markedly low values in the final analyses (type I/II error of step 3 is 0/0.01), suggesting that our model choice analyses were reliable. Model checking analyses indicated that the data simulated under the chosen model and posteriors fitted well the observed genetic data of the OFM (Fig. S3, Table S8, Supporting information).

In step 1, scenario testing analyses of the relationships among the three southern populations suggested a formation of ND from admixture of CD and SL and a split of CD from SL (Fig. S3, Table S8, Supporting information). In step 2, a split of CE from ND was supported among the six scenarios of southern origin, while the early split of NO from CE rather than the southern populations was supported among the two scenarios of central origin of the OFM (Fig. S3, Table S8, Supporting information). In step 3, the three competing hypotheses of southern, central and northern origin of the OFM were compared, revealing that the southern origin of the OFM was the most likely (Fig. S3, Table S8, Supporting information). Our analyses identified both split and admixture events during the dispersal of the OFM from southern to northern regions of China, corresponding to the nested population structure inferred from BAPS analysis. The analyses

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3 Pairwise F_{ST} values of 21 populations of *Grapholita molesta* based on the combined mitochondrial genes

Population	SL	ND	NB	WH	NJ	CD	TS	YL	ZZ	YA	TG	QD	SJ	BD	DX	MT	PG	YQ	DL	SY	KD
SL	0.0000																				
ND	0.2403	0.0000																			
NB	0.1759	0.2277	0.0000																		
WH	0.5667	0.3068	0.4639	0.0000																	
NJ	0.4990	0.2700	0.3934	0.0654	0.0000																
CD	0.6375	0.6233	0.6059	0.6627	0.6529	0.0000															
TS	0.3254	0.2626	0.2590	0.2661	0.2415	0.3385	0.0000														
YL	0.5590	0.3616	0.4714	0.1588	0.0683	0.6829	0.2907	0.0000													
ZZ	0.3211	0.1695	0.1930	0.1161	0.0484	0.6305	0.2191	0.1143	0.0000												
YA	0.3081	0.1423	0.2115	0.0510	0.0275	0.5842	0.1412	0.0878	**0.0057**	0.0000											
TG	0.3448	0.1392	0.2524	0.1557	0.1182	0.6365	0.2459	0.2358	0.0627	0.0487	0.0000										
QD	0.3809	0.1804	0.2766	0.1633	0.0958	0.6716	0.2803	0.1907	**0.0382**	0.0455	**0.0097**	0.0000									
SJ	0.3576	0.1690	0.2591	0.2375	0.1881	0.6490	0.2808	0.3006	0.1088	0.1048	**0.0064**	**0.0257**	0.0000								
BD	0.3815	0.2275	0.3109	0.4571	0.3994	0.6498	0.3379	0.4963	0.2811	0.2677	0.1554	0.1929	0.0845	0.0000							
DX	0.3754	0.2147	0.2924	0.3940	0.3346	0.6638	0.3378	0.4336	0.2221	0.2216	0.0952	0.1212	0.0323	**0.0227**	0.0000						
MT	0.3947	0.1988	0.2957	0.2464	0.1842	0.6593	0.2879	0.2890	0.1077	0.1063	**0.0136**	**0.0190**	**-0.0018**	0.1296	0.0467	0.0000					
PG	0.2857	0.1249	0.2025	0.2371	0.1828	0.6400	0.2679	0.2791	0.0849	0.0936	**0.0179**	**0.0271**	**0.0022**	0.0645	0.0327	**0.0178**	0.0000				
YQ	0.4600	0.2901	0.3607	0.4181	0.3457	0.6683	0.3433	0.4504	0.2418	0.2409	0.1424	0.1546	0.0681	0.1042	0.0536	0.0837	0.0786	0.0000			
DL	0.2875	0.1602	0.2451	0.3432	0.2909	0.6475	0.3056	0.3867	0.1836	0.1830	0.0934	0.1270	0.0711	0.0349	0.0596	0.1042	**0.0320**	0.1309	0.0000		
SY	0.3514	0.1854	0.2306	0.3464	0.2820	0.6348	0.2769	0.3917	0.1480	0.1507	**0.0523**	0.0720	**0.0039**	0.0485	**0.0084**	**0.0251**	**-0.0009**	0.0512	0.0553	0.0000	
KD	0.4917	0.2536	0.3600	0.1644	0.0580	0.6097	0.1828	0.1188	0.0669	**0.0351**	0.1357	0.1244	0.2018	0.3955	0.3399	0.2023	0.1862	0.3668	0.2895	0.2886	0.0000

The bold numbers indicate no significant difference between the two populations (P = 0.05).

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

13652944, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

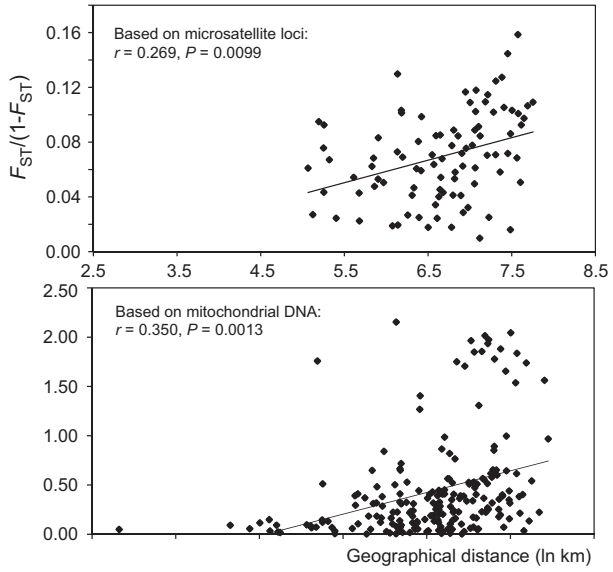


Fig. 4 Scatter plots of genetic distance vs. geographical distance for pairwise population comparisons (both analyses are calculated from 100 000 randomizations).

of microsatellite loci also revealed a southern origin of the OFM and a dispersal routes from southern to the northern region of China. The time for the emergence of CD, ND, CE and NO was estimated to be 268.0 (5% and 95% quantiles: 124.0, 437.0), 54.2 (28.0, 77.2), 32.6 (13.6, 48.0) and 25.1 (6.8, 42.4) Ka, respectively (Fig. S3, Supporting information).

Demographic history

Mismatch distribution analyses uncovered a multimodal distribution for all groups (Fig. 6). However, a sudden expansion model with small and insignificant Harpending's raggedness (HR) index could not be rejected (HR = 0.011–0.164, P = 0.103–0.870), demonstrated a demographic expansion (Rogers & Harpending 1992) or a range expansion recently (Ray *et al.* 2003; Excoffier 2004).

The effective population size estimated for each group revealed different profiles of historical demography. The population size of the SL, ND and major lineage of CD remained stable. The minor lineages of CD and CE exhibited slight decrease in population size, corresponding to their limited distribution in three sampling sites. The major lineage of CE showed slight expansion of population size in the past recent years. The obvious increase in population size was found in the recently derived group of NO at about 5 Ka, demonstrated demographic and range expansion, together revealed by the mismatch distribution analysis (Fig. 6).

Discussion

Origin of the OFM

The ancestral populations usually possess higher genetic diversity than the derived populations (Savolainen *et al.* 2002; Ma *et al.* 2012; Zhang *et al.* 2012). High genetic diversity in both mitochondrial genes and microsatellite loci was found in the studied populations of the OFM, supporting the assumption that China is the native range of this moth (Rothschild & Vickers 1991; Kirk *et al.* 2013). The scenario tests based on ABC method revealed that the southern area is the most likely origin of the OFM in China. This origin is congruent with the high genetic discontinuity (Fig. 5) and high proportion of private haplotype (Fig. S1, Supporting information) in southern populations. Further analysis among three southern populations indicated that SL was the most ancestral one, sharing no haplotype with the other populations (Fig. S1, Supporting information).

In the population genetic structure analysis, the SL and KD clustered to the same group in GENELAND analysis based on the mitochondrial genes. However, most analysis such as the BAPS and SPLITSTREE analysis indicated that these two populations belong to different clusters (Fig. 1) and haplotype lineages (in black and dark red in Fig. 3). It is probably that the South Korean (KD) population is derived from China. This is congruent with the result of the global study of the OFM that the number of private alleles was low in East Asian population from Japan, compared to the mainland Asia (Kirk *et al.* 2013). The gathering of the SL and KD in GENELAND analysis might be caused by the high proportion of private haplotypes and extremely low proportion of the most common haplotypes (Fig. S1, Supporting information).

Quaternary glaciations are often interpreted as a major factor in shaping the biodiversity of the most extant species (Hewitt 2004). Based on the gene tree, the estimated divergence time between the two haplotype lineages is 254 Ka, which falls within the Pleistocene (2588 to 12 Ka). Despite having temperatures similar to those of glaciated areas in North America and Europe, China remained unglaciated, except at higher elevations (Shi *et al.* 2011). Therefore, the persistence of the OFM is possible in the Shilin (SL) area. This area, having a subtropical climate, either was found to be a refuge during the Quaternary climatic change or is the place of origin of many species (Yang *et al.* 2008).

Host plant dispersal usually plays an important role in passively invading insects. Although many host plants for the OFM were reported, the peach appears to be its original host (Rothschild & Vickers 1991). Peach

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

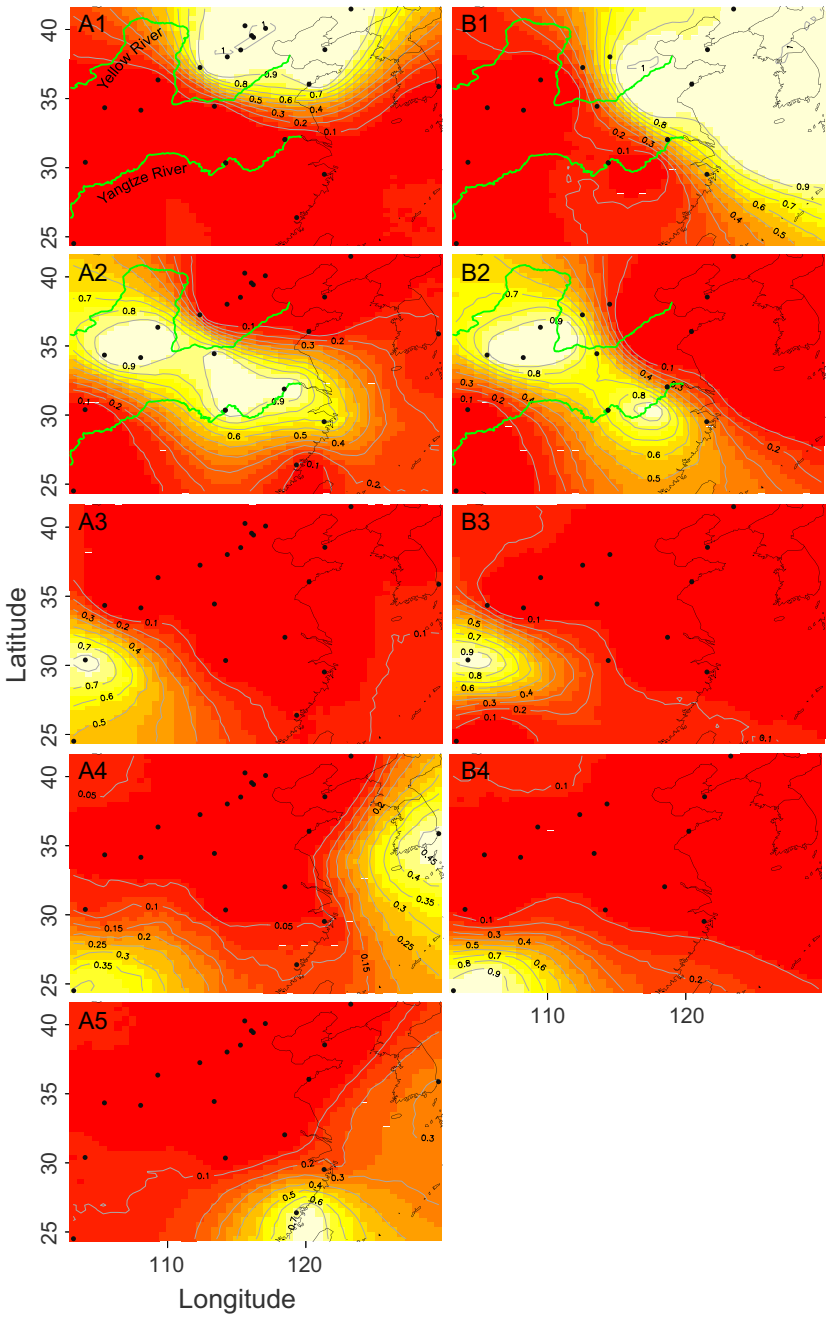


Fig. 5 Maps of the posterior probabilities of population membership inferred by GENELAND. The combined mitochondrial genes inferred five clusters (A1–A5), while the microsatellite loci inferred four clusters (B1–B4). Contour lines indicate the spatial position of the genetic discontinuities. Lighter shading indicates higher probabilities of population membership. Reduced number of 14 populations was used in the analysis based on the microsatellite loci.

orchards were usually selected in their newly introduced range (Quaintance & Wood 1916; Blomefield & Barnes 2000), which enhanced the developmental and survival rate of the larvae by providing abundant shoots in the earlier summer (Myers *et al.* 2007). Current research has clearly demonstrated that China is the native habitat of the peach and that the southwest region was the native location of the ancestral species *Prunus mira* Koehne of the modern peach varieties distributed in southeastern Tibet (Wang & Zhuang 2001; Faust & Timon 2010; Cao *et al.* 2014). Our inference on the origin of the OFM agrees with the origin of its native host plant in China.

Dispersal of the OFM in China

The use of ABC method can be complex and risky for the parameter setting, scenario definition and population sampling. In our analyses, the input populations were defined based on population genetic structure of the OFM, which was recommended as complementarity of ABC method (Lombaert *et al.* 2014). We also

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

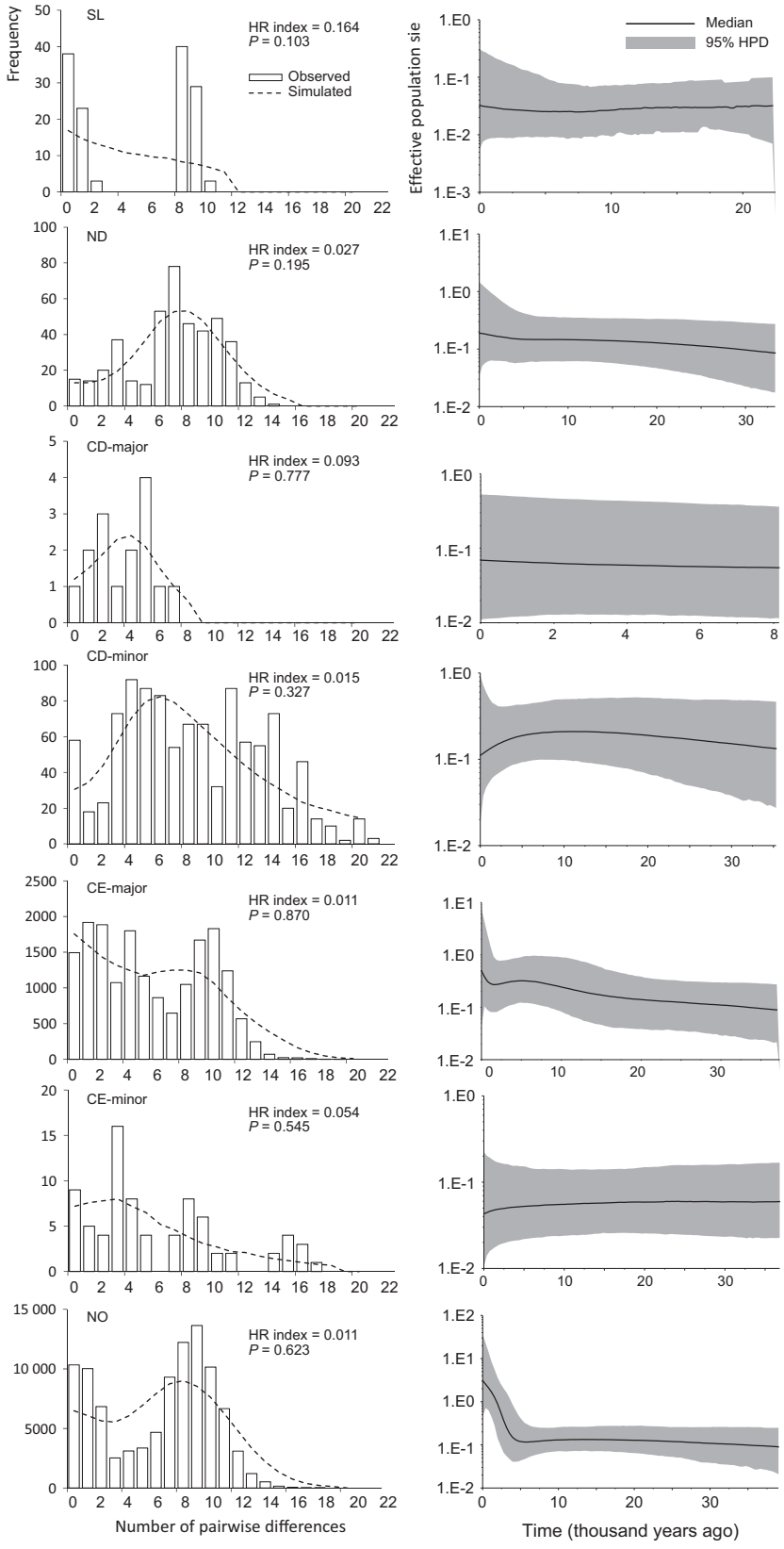


Fig. 6 Mismatch distributions (left) and Bayesian skyline plots (right) of five population groups of *Grapholita molesta* identified by GENELAND. The two haplotype lineages identified by SPLITSTREE in CD and CE groups were analysed separately. The CD, SL and ND were calculated based on the sampled populations of CD, SL and ND, respectively. The CE was calculated based on the sampled populations of TS, YL, YA and WH. The NO was calculated based on the sampled populations of NJ, NJ, ZZ, TG, SJ, QD, DL and SY. Harpending's raggedness (HR) indices are shown in the mismatch distribution. For Bayesian skyline plots, the middle lines represent the median estimates of the effective population size, and the shaded areas represent 95% highest posterior densities (95% HPD). Abbreviations of populations and groups are described in Table 1.

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

considered the variation and fixation of population size in scenario definition. The mitochondrial genes and microsatellite loci were analysed separately, showing consistent results in scenario selection.

The ABC analyses indicated that the CD population splits from SL at an estimated time of 268 Ka, which is coincident with the Penultimate glaciation, MIS6 (marine isotope stages) to MIS8 (before 130 thousand years BP), in China (Shi *et al.* 2011). Afterwards, the OFM dispersed to ND from SL with an admixture of CD population, at an estimated time of 54.2 Ka and then further dispersed northward to central and northern China at 32.6 and 25.1 Ka, respectively, before the last glacial maximum (LGM) (18–25 Ka) (Clark *et al.* 2009; Shi *et al.* 2011), in China. Unlike species in temperate North America and Europe which commonly expanded in the post-LGM era, our result indicates that the spatial expansion of the OFM occurred within the Pleistocene, at which time, China remained unglaciated, except at higher elevations (Shi *et al.* 2011). This case is similar to that found in the striped rice stem borer *Chilo suppressalis* (Crambidae) (Meng *et al.* 2008) and the shrub *Sophora davidii* (Leguminosae) (Fan *et al.* 2013). The finding of the plant species *P. mira* in the Muli area of Sichuan Province (Yang 1985) confirms a possible northward dispersal of the OFM associated with its host plant.

The original and dispersed populations of the OFM showed nested and differentiated structure with a pattern of isolation by distance, forming three southern, and the central and northern groups. This genetic pattern is concordant with geographical distribution. The southern populations are more differentiated and genetically diversified than the northern populations. This phenomenon suggests that the high elevation in southern China interspersed with many mountain ranges probably retarded the early dispersal and gene flow of the OFM, although we could not exclude the fact that the original area tends to persist high genetic diversity. Significant geographical structure and high genetic diversity in southern China have been detected in populations of *C. suppressalis* (Meng *et al.* 2008), the mosquito *Anopheles jeyporiensis* (Chen *et al.* 2004) and the Silver Pheasant *Lophura nycthemera* (Dong *et al.* 2013). Although the genetic structure of the central and northern populations might be diminished by accelerated rate of dispersal and human-mediated admixture, two geographical and genetic groups were identified, as featured in SPLITSTREE (Fig. S1, Supporting information, Hap16 and Hap 27 in Fig. 3), BAPS (red and yellow clusters in Fig. 1) and GENELAND (CE and NO group) analyses. The boundaries of two groups were coincident with the two great rivers run through China, that is the Yangtze River and the Yellow River (Figs 1 and 5), suggesting that further dispersal of the OFM in central and northern China was probably influenced by river ranges. River range has been detected as a substantial barrier to gene flow in China in species of insects (Zhu *et al.* 2005; Meng *et al.* 2008), plants (Zhang & Sun 2011) and even birds (Li *et al.* 2009; Song *et al.* 2009).

Demographic history of the OFM

From the view of genetic diversity, the SL population exhibited a relatively low genetic diversity, while the CD population exhibited the highest genetic diversity among the sampled populations (Table 2). High genetic diversity out of the ancestral population was reported, which is likely a consequence of the admixture of divergent lineages colonizing the continent from separate refugia (Petit *et al.* 2003). It is also likely that during a range change, a retreating rear edge will suffer shrinkage, dissection and extinction, so that the last surviving population should be severely bottlenecked (Hewitt 2000). We presumed that both the effect of climate oscillation and the limited host plants in the SL area, which is near the glaciated Qinghai–Tibetan Plateau, might cause the reduction in the effective population size.

The DIYABC and molecular dating analyses indicated the dispersal of OFM to its currently distributed area before the growth of the ice sheets to their maximum positions in China (Shi *et al.* 2011). Although no massive ice sheet was developed in central China during glacial periods, the tremendous global climatic changes might affect the distribution of the OFM, thus forming multiple refugia, as has been demonstrated in China in other insects (Meng *et al.* 2008; Zhang *et al.* 2009), birds (Qu *et al.* 2012; Dong *et al.* 2013) and plants (Wang *et al.* 2008; Zhou *et al.* 2010; Li *et al.* 2012; Fan *et al.* 2013). The multimodal mismatch distribution in all groups (Fig. 6) and the nested population structure revealed by BAPS analysis (Fig. 1) might indicate the existence of multiple refugia in China. The cluster present in CD, TS and YL populations (blue in Fig. 1) identified by BAPS analysis is likely be originated from a refuge near the CD area. This area belonging to the Sichuan Basin has a humid subtropical monsoonal climate. The limited distribution of this cluster might be due to local adaption to the environment. The secondary dispersal of the individuals from SL possibly leads to the existence of two haplotype lineages, thus increased the genetic diversity of the CD and TS populations.

The expansion of population size in NO occurred in the last 5000 years identified by Bayesian skyline plot analysis (Fig. 6), after the Quaternary glaciations. The ancestral species of peach was recorded to be disseminated to the north and east of its native region by the

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ancient Qiang people several thousand years ago (Wang & Zhuang 2001). Archaeological records indicate that the peach was eaten by humans approximately 7000 years ago (Cao *et al.* 2014). The recent expansion of OFM might be associated with its host plant dissemination in the history of human activities, which is an important factor leading to the current genetic structure of many species (Meng *et al.* 2008; Dsouli-Aymes *et al.* 2011; Ma *et al.* 2012).

Conclusion

This study widely examined the population genetic structure of the OFM in its presumed native range. Using the ABC method, the exact range of origin was traced to southern China near the Shilin area of Yunnan Province, which has been proved to be the place of origin of many species. This inferred area of origin of the OFM is congruent with the origin of *P. mira*, the oldest progenitor of the modern peach species. The early divergence and dispersal of the OFM in southern region of China was dated to the Penultimate glaciation of Pleistocene. Further dispersal from southern to northern region of China occurred before the LGM. The expansion of population size in the derived populations in northern region of China occurred after the LGM, which might be related to human activities on the cultivation of its host plant. Our results indicated that the current distribution and structure of the OFM were complicatedly influenced by climatic and geological events and human activities of cultivation and wide dissemination of peach in ancient China. We provide a successful example of uncovering the genetic structure and its underlying factors of an agricultural pest insect in East Asia. The findings in our study might also shed light on the research studies of peach planting and Quaternary glaciations in the area.

Acknowledgements

We thank the anonymous reviewers and the Subject Editor Prof. Graham Stone for their valuable comments, Dr. Julien Foucaud from the Centre de Biologie pour la Gestion des Populations of the French National Institute for Agricultural Research, Dr. Cheng-Min Shi and Prof. De-Xing Zhang from the Institute of Zoology of the Chinese Academy of Sciences and Prof. Li-Jun He from the East China Normal University for their help on the data analysis. We also thank those who contributed to the sampling of the specimens for this research, including Zong-Jiang Kang, Liang Zhu, Hong Lu, Guang-Hang Qiao, Fei Yu, Rui Guo, Ming Xie, Jing Li, Yuan-Xi Li, Li-Ping Hu, Geng-Rui Zhu, Rui-De Jiang, Zhi-Guo Zhao, Hu Chen, Chun-Hai Guan and Long-Long Weng. The research was funded by the National Basic Research Program of China (2013CB127600), the National Natural Science Foundation of China (31101661, 31472025), the Chinese Agriculture Research System (CARS-31), the Special Fund for Agro-scientific Research in the Public Interest (201103024) and the Innovation Foundation of IPEP-BAAFS (CXJJ2012A03).

References

Baele G, Lemey P, Bedford T *et al.* (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.

Baele G, Li WLS, Drummond AJ, Suchard MA, Lemey P (2013) Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, **30**, 239–243.

Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Blomefield T, Barnes B (2000) Integrated control of oriental fruit moth, *Grapholita molesta*, on peaches using a spray-date prediction model. *Acta Horticulturae*, (ISHS) **525**, 161–168.

Bonnet E, Van de Peer Y (2002) ZT: a software tool for simple and partial Mantel tests. *Journal of Statistical software*, **7**, 1–12.

Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 193–216.

Cao K, Zheng Z, Wang L *et al.* (2014) Comparative population genomics reveals the domestication history of the peach, *Prunus persica*, and human influences on perennial fruit crops. *Genome Biology*, **15**, 415.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Chen B, Harbach RE, Butlin RK (2004) Genetic variation and population structure of the mosquito *Anopheles jeyporiensis* in southern China. *Molecular Ecology*, **13**, 3051–3056.

Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J (2013) Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Molecular Biology and Evolution*, **30**, 1224–1228.

Clark PU, Dyke AS, Shakun JD *et al.* (2009) The last glacial maximum. *Science*, **325**, 710–714.

Cornuet J-M, Pudlo P, Veyssier J *et al.* (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187–1189.

Cory EN (1925) *Laspeyresia molesta* Busck as a quince pest. *Journal of Economic Entomology*, **18**, 199–202.

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Dong L, Heckel G, Liang W, Zhang Y (2013) Phylogeography of Silver Pheasant (*Lophura nycthemera* L.) across China: aggregate effects of refugia, introgression and riverine barriers. *Molecular Ecology*, **22**, 3376–3390.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Drummond A, Rambaut A, Shapiro B, Pybus O (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

Dsouli-Aymes N, Michaux J, De Stordeur E *et al.* (2011) Global population structure of the stable fly (*Stomoxys calcitrans*) inferred by mitochondrial and nuclear sequence data. *Infection, Genetics and Evolution*, **11**, 334–342.

Endersby NM, McKechnie SW, Ridland PM, Weeks AR (2006) Microsatellites reveal a lack of structure in Australian populations of the diamondback moth, *Plutella xylostella* (L.). *Molecular Ecology*, **15**, 107–118.

Estoup A, Lombaert E, Marin JM *et al.* (2012) Estimation of demo-genetic model probabilities with approximate Bayesian computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources*, **12**, 846–855.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853–864.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Fagundes NJ, Ray N, Beaumont M *et al.* (2007) Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences, USA*, **104**, 17614–17619.

Fan DM, Yue JP, Nie ZL *et al.* (2013) Phylogeography of Sophora davidii (Leguminosae) across the 'Tanaka-Kaiyong Line', an important phytogeographic boundary in Southwest China. *Molecular Ecology*, **22**, 4270–4288.

Faust M, Timon B (2010) Origin and dissemination of peach. *Horticultural Reviews*, **17**, 331–379.

Gong YJ, Shi BC, Kang ZJ, Zhang F, Wei SJ (2012) The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck)(Lepidoptera: Tortricidae). *Molecular Biology Reports*, **39**, 2893–2900.

Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Guillot G, Santos F, Estoup A (2008) Analysing georeferenced population genetics data with GENELAND: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, **24**, 1406–1407.

Guindon S, Dufayard JF, Lefort V *et al.* (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PHYML 3.0. *Systematic Biology*, **59**, 307–321.

Halpern BS, Walbridge S, Selkoe KA *et al.* (2008) A global map of human impact on marine ecosystems. *Science*, **319**, 948–952.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt G (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences*, **359**, 183–195.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Kerdelhué C, Zane L, Simonato M *et al.* (2009) Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**, 220.

Kirk H, Dorn S, Mazzi D (2013) Worldwide population genetic structure of the oriental fruit moth (*Grapholita molesta*), a globally invasive pest. *BMC Ecology*, **13**, 12.

Kobayashi T, Sakurai T, Sakakibara M, Watanabe T (2011) Multiple origins of outbreak populations of a native insect pest in an agro-ecosystem. *Bulletin of Entomological Research*, **101**, 313–324.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PARTITIONFINDER: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

Li SH, Yeung CKL, Feinstein J *et al.* (2009) Sailing through the Late Pleistocene: unusual historical demography of an East Asian endemic, the Chinese Hwamei (*Leucodioptron canorum canorum*), during the last glacial period. *Molecular Ecology*, **18**, 622–633.

Li Y, Yan HF, Ge XJ (2012) Phylogeographic analysis and environmental niche modeling of widespread shrub *Rhododendron simsii* in China reveals multiple glacial refugia during the last glacial maximum. *Journal of Systematics and Evolution*, **50**, 362–373.

Lombaert E, Guillemaud T, Lundgren J *et al.* (2014) Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, **23**, 5979–5997.

Lyons JI, Pierce AA, Barribeau SM *et al.* (2012) Lack of genetic differentiation between monarch butterflies with divergent migration destinations. *Molecular Ecology*, **21**, 3433–3444.

Ma C, Yang P, Jiang F *et al.* (2012) Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Molecular Ecology*, **21**, 4344–4358.

Meng XF, Shi M, Chen XX (2008) Population genetic structure of *Chilo suppressalis* (Walker)(Lepidoptera: Crambidae): strong subdivision in China inferred from microsatellite markers and mtDNA gene sequences. *Molecular Ecology*, **17**, 2880–2897.

Myers CT, Hull LA, Krawczyk G (2007) Effects of orchard host plants (apple and peach) on development of oriental fruit moth (Lepidoptera: Tortricidae). *Journal of Economic Entomology*, **100**, 421–430.

van Oosterhout C, Hutchinson WF, WILLS DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology and Evolution*, **27**, 1659–1672.

Parker M, Thompson JN, Weller SG (2001) The population biology of invasive species. *Annual Review of Ecology, Evolution, and Systematics*, **32**, 305–332.

Petit RJ, Aguinagalde I, de Beaulieu J-L *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Pinero JC, Dorn S (2009) Response of female oriental fruit moth to volatiles from apple and peach trees at three phenological stages. *Entomologia Experimentalis et Applicata*, **131**, 67–74.

Porretta D, Canestrelli D, Bellini R, Celli G, Urbanelli S (2007) Improving insect pest management through population genetic data: a case study of the mosquito *Ochlerotatus caspius* (Pallas). *Journal of Applied Ecology*, **44**, 682–691.

Qu Y, Zhang R, Quan Q *et al.* (2012) Incomplete lineage sorting or secondary admixture: disentangling historical diver-

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

gence from recent gene flow in the Vinous-throated parrotbill (*Paradoxornis webbianus*). *Molecular Ecology*, **21**, 6117–6133.
Quaintance AL, Wood WB (1916) *Laspeyresia molesta*, an important new insect enemy of the peach. *Journal of Agricultural Research*, **7**, 373–378.
Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, **20**, 76–86.
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.
Rothschild G, Vickers R (1991) Biology, ecology and control of the oriental fruit moth. In: *World Crop Pests. Tortricid Pests: Their Biology, Natural Enemies and Control* (eds van derGeest L, Evenhuis H), pp. 389–412. Elsevier, Amsterdam.
Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T (2002) Genetic evidence for an East Asian origin of domestic dogs. *Science*, **298**, 1610–1613.
Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.
Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.
Shi YF, Zhao JD, Wang J (2011) *New Understanding of Quaternary Glaciations in China*. Shanghai Popular Science Press, Shanghai.
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.
Song G, Qu Y, Yin Z *et al.* (2009) Phylogeography of the *Alcippe morrisonia* (Aves: Timaliidae): long population history beyond late Pleistocene glaciations. *BMC Evolutionary Biology*, **9**, 143.
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.
Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.
Timm AE, Geertsema H, Warnich L (2008) Population genetic structure of *Grapholita molesta* (Lepidoptera: Tortricidae) in South Africa. *Annals of the Entomological Society of America*, **101**, 197–203.
Torriani MV, Mazzi D, Hein S, Dorn S (2010) Structured populations of the oriental fruit moth in an agricultural ecosystem. *Molecular Ecology*, **19**, 2651–2660.
Wang ZH, Zhuang EJ (2001) *China Fruit Monograph: Peach Flora*. China Forestry Publishing House, China, Beijing.
Wang FY, Gong X, Hu CM, Hao G (2008) Phylogeography of an alpine species *Primula secundiflora* inferred from the chloroplast DNA sequence variation. *Journal of Systematics and Evolution*, **46**, 13–22.
Wei SJ, Shi BC, Gong YJ *et al.* (2013) Genetic structure and demographic history reveal migration of the diamondback moth *Plutella xylostella* (Lepidoptera: Plutellidae) from the southern to northern regions of China. *PLoS ONE*, **8**, e59654.
Yang SJ (1985) The discovery of the smooth pit peach *Prunus mira* Koehne in Tibetan Autonomous County of Muli of Sichuan Province. *Journal of Fruit Science*, **6**, 120. (in Chinese).
Yang F, Li YF, Ding X, Wang XQ (2008) Extensive population expansion of Pedicularis longiflora (Orobanchaceae) on the Qinghai-Tibetan Plateau and its correlation with the Quaternary climate change. *Molecular Ecology*, **17**, 5135–5145.
Yang XM, Sun JT, Xue XF, Li JB, Hong XY (2012) Invasion genetics of the western flower thrips in China: evidence for genetic bottleneck, hybridization and bridgehead effect. *PLoS ONE*, **7**, e34567.
Yetter W, Steiner L (1932) Efficiency of bait traps for the oriental fruit moth as indicated by the release and capture of marked adults. *Journal of Economic Entomology*, **25**, 106–116.
Zhang T, Sun H (2011) Phylogeographic structure of *Terminalia franchetii* (combretaceae) in southwest China and its implications for drainage geological history. *Journal of Plant Research*, **124**, 63–73.
Zhang H, Yan J, Zhang G, Zhou K (2008) Phylogeography and demographic history of Chinese black-spotted frog populations (*Pelophylax nigromaculata*): evidence for independent refugia expansion and secondary contact. *BMC Evolutionary Biology*, **8**, 21.
Zhang DX, Yan LN, Ji YJ, Hewitt GM, Huang ZS (2009) Unexpected relationships of substructured populations in Chinese *Locusta migratoria*. *BMC Evolutionary Biology*, **9**, 144.
Zhang B, Edwards O, Kang L, Fuller S (2012) Russian wheat aphids (*Diuraphis noxia*) in China: native range expansion or recent introduction? *Molecular Ecology*, **21**, 2130–2144.
Zheng Y, Peng X, Liu G *et al.* (2013) High genetic diversity and structured populations of the oriental fruit moth in its range of origin. *PLoS ONE*, **8**, e78476.
Zhou TH, Li S, Qian ZQ *et al.* (2010) Strong phylogeographic pattern of cpDNA variation reveals multiple glacial refugia for *Saruma henryi* Oliv. (Aristolochiaceae), an endangered herb endemic to China. *Molecular Phylogenetics and Evolution*, **57**, 176–188.
Zhu ZH, Ye H, Zhang Z (2005) Genetic relationships among four *Bactrocera cucurbitae* geographic populations in Yunnan Province. *Chinese Journal of Applied Ecology*, **16**, 1889–1892.
Zlojutro M, Rubicz R, Devor EJ *et al.* (2006) Genetic structure of the Aleuts and Circumpolar populations based on mitochondrial DNA sequences: a synthesis. *American Journal of Physical Anthropology*, **129**, 446–464.

S.J.W. conceived and designed the research and wrote the manuscript. S.J.W. and L.J.C. performed the DNA sequencing and the microsatellite loci determination. S.J.W. analysed the data. S.J.W., F.Z. and X.J.G. organized the collection of the specimens. S.J.W., X.X.C., Y.J.G., S.W., B.C.S. and Y.M.W. discussed the results. All authors read and approved the final manuscript.

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Data accessibility

Data and scripts are deposited in Dryad (doi:10.5061/dryad.k1m1g), and mitochondrial DNA haplotypes are deposited in GenBank (KF163966-KF164208, KT004678-KT004683).

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Network analysis of the haplotype relationships.

Appendix S2 Phylogenetic analysis of the haplotypes.

Appendix S3 Splitstree from each of the seven mitochondrial genes.

Appendix S4 Multivariate analyses of microsatellite data using discriminant analysis of principal components (DAPC).

Appendix S5 Population structure analysis of the microsatellite loci using STRUCTURE.

Appendix S6 DIYABC analyses based on the microsatellite loci.

Table S1 Discovery of *Grapholita molesta* in different countries.

Table S2 Amplification and sequencing primers for the mitochondrial gene sequences used in this study.

Table S3 Characteristics of eight polymorphic microsatellite loci used in this study.

Table S4 Range of priors for the parameters of the scenarios in the DIYABC analyses using the mitochondrial genes.

Table S5 Parameters of genetic diversity and demographic analysis of seven mitochondrial genes.

Table S6 Summary statistics of the microsatellite loci examined in the 14 populations.

Table S7 Pairwise F_{ST} values of 14 populations based on the microsatellite loci.

Table S8 Scenario selection, evaluation of confidence and model checking in scenario selection for the DIYABC analyses using the mitochondrial genes.

Fig. S1 Percentage of total and private haplotypes in each population.

Fig. S2 Scatter plots of genetic distance vs. geographical distance for pairwise population comparisons when the minor lineage was excluded.

Fig. S3 Results of DIYABC analyses using mitochondrial genes.

© 2015 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

1365294x, 2015, 16, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.13300 by CochraneChina, Wiley Online Library on [25/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License