

SCIENTIFIC REPORTS

OPEN

Population genetic structure of *Diaphorina citri* Kuwayama (Hemiptera: Liviidae): host-driven genetic differentiation in China

Lixue Meng^{1,2}, Yongmo Wang², Wen-Hua Wei³ & Hongyu Zhang^{1,2}

The Asian citrus psyllid *Diaphorina citri* Kuwayama is a major pest in citrus production, transmitting *Candidatus Liberibacter asiaticus*. It has spread widely across eastern and southern China. Unfortunately, little is known about the genetic diversity and population structure of *D. citri*, making pest control difficult. In this study, nine specifically developed SSR markers and three known mitochondrial DNA were used for population genetics study of *D. citri* using 225 samples collected from all 7 distribution regions in China. Based on the SSR data, *D. citri* was found highly diverse with a mean observed heterozygosity of 0.50, and three subgroups were structured by host plant: (i) Shatangju, NF mandarin and Ponkan; (ii) Murraya paniculata and Lemon; (iii) Citrus unshiu, Bingtangcheng, Summer orange and Navel. No significant genetic differences were found with mtDNA data. We suggested the host-associated divergence is likely to have occurred very recently. A unimodal distribution of paired differences, the negative and significant Tajima's *D* and Fu's *F_s* parameters among mtDNA suggested a recent demographic expansion. The extensive citrus cultivation and increased suitable living habitat was recommended as a key for this expansion event.

The Asian citrus psyllid, *Diaphorina citri* Kuwayama (Hemiptera: Liviidae), is the most serious pest of citrus worldwide, primarily due to its role as a vector for "*Candidatus Liberibacter asiaticus*", the bacterium that causes the highly destructive Asian huanglongbing (HLB) (citrus greening)¹. HLB affects all citrus cultivars^{2,3}. Severely HLB-infected trees often appear stunted, sparsely foliated and die back, losing their economic viability. Ledford⁴ reported that this disease has slashed US orange production in half over the past decade, and threatens to destroy the US\$3.3-billion industry entirely. To date there is no effective control technique for this disease⁵, and control of *D. citri* has been one of three critical components of HLB management in addition to planting pathogen-free nursery stock and removing inoculum by destroying infected trees⁶.

Diaphorina citri was first described in Taiwan in 1907 and currently has spread throughout much of Asia^{6,7}. In Brazil, *D. citri* was found in the 1940s⁸, expanded its range to Florida in the late 1990s⁹, and is now well established in citrus producing regions of the United States^{10,11} as well as Mexico, Belize, Costa Rica, the Caribbean and South America¹². The wide distribution of *D. citri* and probable high genetic diversity make any suppression efforts challenging. Taking China as an example, after a first found in Taiwan in 1907⁶, *D. citri* spread quickly to the coastal provinces including Guangdong, Fujian and Guangxi, and then to inland provinces such as Hunan, Jiangxi, Yunnan and Sichuan¹³. During the invasion course, *D. citri* has adapted itself to a wide range of geographic environments and host species, becoming ubiquitous in almost all citrus growing areas in China^{14,15}.

Deciphering genetic diversity, population structure, as well as history and adaptation processes is essential for developing effective suppression protocols and sustainable management strategies for insect pests¹⁶. DNA markers such as mtDNA, SSRs, RAPD, AFLP and ESTs have been used as popular marker systems in insect genetics research¹⁷. In the last few years, much attention has been drawn to phylogeographic studies of *D. citri* populations;

¹Key Laboratory of Horticultural Plant Biology (MOE), State Key Laboratory of Agricultural Microbiology, Institute of Urban and Horticultural Entomology, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, 430070, China. ²Hubei Insect Resources Utilization and Sustainable Pest Management Key Laboratory, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, Hubei, P.R. China. ³Department of Women's and Children's Health, Dunedin School of Medicine, University of Otago, Dunedin, 9016, New Zealand. Correspondence and requests for materials should be addressed to H.Z. (email: hongyu.zhang@mail.hzau.edu.cn)

Locus ID	Primer sequences (5'–3')	Repeat motif	Size range (bp) ^a	Ta (°C)	Accession no.	A	Ho	He	F _{IS}
Dci-1	F:GTTGGGAATAGGCCTGACAAR:ATCAGGACGGTGGACACTTT	(TC) ₂₂	166–218	61	KY053427	6	0.636	0.483	–0.331
Dci-7	F:GATGGAAGAGAGAAAAGGGAGAR:GGGTGGGGTCATCAAATA	(AG) ₂₀	217–339	60	KY053428	13	0.638	0.798	0.187
Dci-8	F:TAACCATCGTCAGGCAATCAR:TGGGAGGGAAAAGATTGAGAA	(TC) ₁₅	250–276	62	KY053429	12	0.197	0.837	0.763
Dci-22	F:GCTTCTTGTCGCTTGTGCR:GAGTTTGGTGGATGTGCC	(GA) ₂₇	128–216	60	KY053430	18	0.915	0.904	–0.015
Dci-26	F:TTAGTGGAGATGCCTGATR:TCGCCAACATACGAAACG	(TTC) ₂₅	104–110	60	KY053431	3	0.877	0.571	–0.546
Dci-29	F:CAATCTTGATACTGTGGGTR:TGGAAGTGAGGGTGAGAA	(TTC) ₂₂	184–262	60	KY053432	12	0.192	0.317	0.387
Dci-33	F: TGCTTACTTTGTAGACCR:ATACCTCCATAACATCCC	(GA) ₈	125–159	60	KY053433	13	0.102	0.234	0.560
Dci-34	F:CCTGGGATAATGGATGACR:CTGTGTACCCTGCTTCG	(GT) ₇	104–122	60	KY053434	11	0.442	0.442	–0.016
Dci-36	F:TATCAATCCTACCAAAGTR:TGAAGGGGCATTACAACA	(GA) ₉	189–209	60	KY053435	11	0.479	0.727	0.338

Table 1. Characteristics of nine polymorphic SSR primers developed in *D. citri*. A, number of alleles detected; Ta, annealing temperature using the genotyping protocol of Schuelke¹¹⁴. H_o , observed heterozygosity; H_e , expected heterozygosity; ^athe allele sizes of the PCR products including the length of the fluorescent-labeling M13 (–21) sequencing primer (18 bp); F_{IS} , inbreeding coefficient; *Significant deviation from Hardy-Weinberg equilibrium ($P < 0.05$).

most of these studies utilized mitochondrial genes (i.e., the COI gene) as genetic markers. Studies of *D. citri* populations with the COI gene in America suggested two independent founder events occurred in South and North America¹⁸. Boykin¹⁹ found two major haplotype groups with preliminary geographic bias between southwestern and southeastern Asia using the COI gene. *Diaphorina citri* was reported to have formed two geographically and genetically structured groups (the eastern versus western side of the state of São Paulo) and has undergone a recent population expansion in Brazil based on COI sequences²⁰.

As a marker, some of the unique characteristics of mtDNA compared to nuclear DNA are a high rate of nucleotide substitution, maternal inheritance, and little or no recombination^{17,21}. Generally, mtDNA is assumed to evolve under neutral or nearly-neutral selection²¹. However, the Neutral theory is only valid for slow evolving genes yet to reach Maximum Genetic Diversity (MGD) and using the fast evolving part is wrong to analyze phylogeny²². Besides, the use of only mtDNA for phylogeographic and population genetic studies is questioned, because within-species diversity does not reflect population size or ecology, and it is difficult to investigate evolutionary questions²³. Utilizing both slow evolving and fast evolving molecular markers together can generate a comprehensive picture of both ancient and recent population histories and thus becomes necessary^{24,25}. Owing to their rapid evolution, extensive genome distribution and high level of polymorphism, as well as apparent neutrality and co-dominant behavior, Short Simple Repeats (SSR) or microsatellite markers have become a useful marker in studies of ecology, evolution, conservation, and population biology of many insect species^{26,27}. Compared to the AFLP system, SSR markers are also sufficiently abundant to generate large numbers of markers¹⁷ and have been widely used in plants, animals, and microorganisms^{28–32}. The joint application of mtDNA and SSR markers in the same population will make the results of phylogeography and population genetics analyses more validated and additional genetic information will be gained²². To date, no study has been undertaken in *D. citri* to investigate the genetic diversity and relationships among populations using different classes of genetic markers simultaneously on the same samples. And, no researcher has tested if mtDNA genes are neutral before using them.

This study aimed at (1) investigating the genetic diversity of *D. citri* populations in China using three mtDNA genes (COI, Cytb and ND5) and nine newly developed SSR markers; (2) investigating the factors shaping the population genetic structure of *D. citri* populations; and (3) investigating the demographic history of *D. citri* populations in China to propose a hypothesis on the expansion of this pest under China conditions. Through the study, we wish to provide a foundation to improve the monitoring and control of *D. citri*.

Results

SSR markers developed. In *D. citri*, 144 SSR loci were obtained after scanning the whole genome with SSR Hunter 1.3. These motifs contained dinucleotide (57.6%) or trinucleotide (38.9%) motifs. The others were less frequent (9.4% tetranucleotide, 1.4% pentanucleotide). Of the 144 potentially amplifiable loci, 50 primer pairs were tested and 20 were selected according to their amplification. Thirteen primer pairs were validated on a panel of samples, i.e., production of specific PCR products of the expected size, and tested for polymorphism. Nine of these 13 loci (69.2%) proved to be polymorphic for all individuals (Table 1). These loci contained 7 dinucleotide and 2 trinucleotide motifs.

All loci showed no evidence of fragment-length homoplasy. A few single nucleotide polymorphisms (SNPs) were identified sporadically in the flanking regions of SSRs in the makers Dci-22 and Dci-26. Since SNPs were not located in the target region of primers and had no effect on the length of amplicons, they would not hamper the amplification of the markers or data analysis³³. Therefore, all nine SSRs were used for further population genetics study.

Overall genetic variation. Different parameters of genetic variation were estimated. The average number of alleles of the identified SSR loci was 11 (ranging from 3 to 18) with mean heterozygosity (H_o) of 0.484 ranging from 0.100 (locus Dci-33) to 0.931 (locus Dci-22) (Table 1). Tests of linkage disequilibrium (LD) between each pair of loci using GENEPOP v3.1 across all populations found no significant linkage suggesting that these SSR

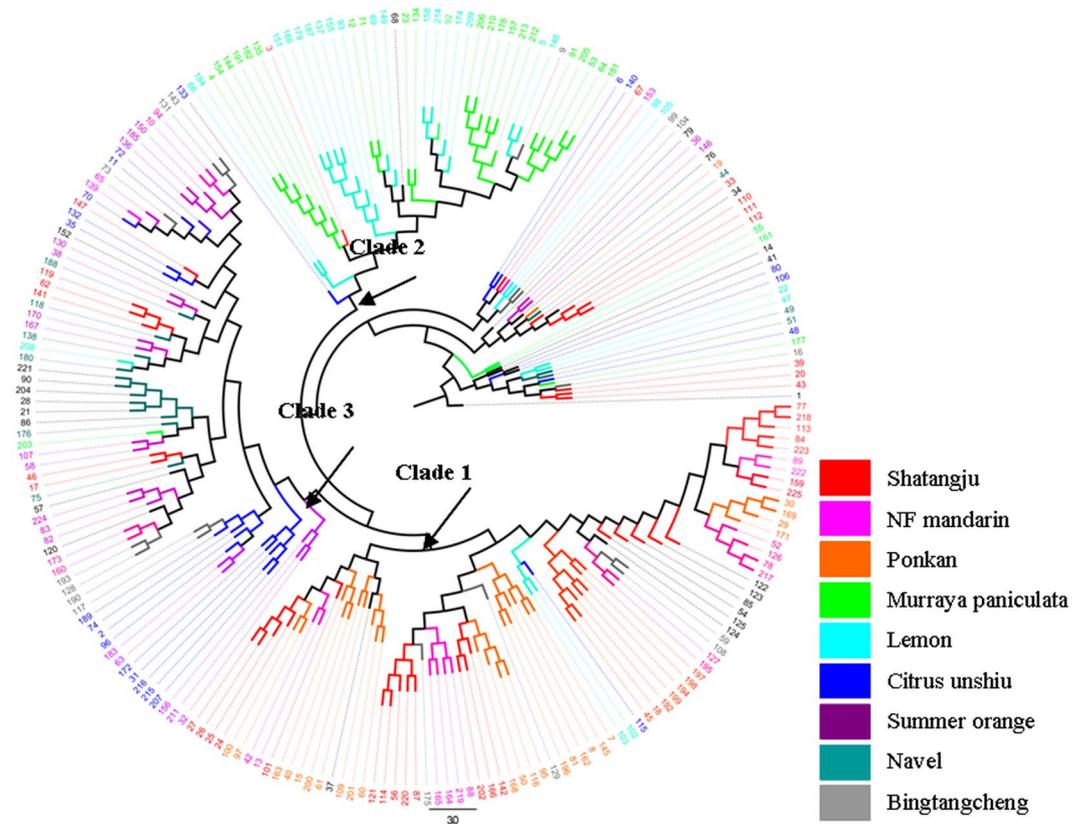


Figure 1. Neighbor-joining tree constructed using pairwise differences for *D. citri* individuals. The populations are color-coded according to the pre-defined ancestry categories.

loci are independent from each other ($p = 0.391$). Four of nine loci were found to have deviated from HWE ($p < 0.05$) after Bonferroni correction in all populations. Further tests (i.e., F_{IS}) revealed that the deviations from HWE may be caused by heterozygote deficiency (Table 1). MICROCHECKER found some evidence of null alleles but each with a low frequency less than 0.081, meaning that population substructure or selection upon the loci may be responsible for the observed HWE deviations. The variation parameters obtained in our work are mostly above the range of those reported SSR for *D. citri* in previous works^{34,35}, suggesting a high variation level in the 9 SSR loci studied.

We sequenced three mitochondrial fragments from 225 *D. citri* individuals. A total of 2398 bp in a concatenated sequence from COI, Cytb and ND5 were analyzed. There were 72 (3%) variable positions in the concatenated sequence (COI had 2.56%, Cytb had 3.34% and ND5 had 3.07%) with a transition: trans-version ratio of 1:1. In total, 26 *D. citri* haplotypes were identified from the concatenated mtDNA in which all populations had low diversity at both the nucleotide (π : 0.00082) and haplotype (Hd: 0.4957) levels.

Phylogeny derived from SSR and mtDNA data. The unrooted NJ phylogram generated from the SSR data indicated a high degree of clustering by host plant species (Fig. 1). Samples from the nine host plants formed three main discrete clusters. *Diaphorina citri* individuals from Shatangju, NF mandarin and Ponkan were clustered in Clade 1; those from *Murraya paniculata* and Lemon in Clade 2; those from Citrus unshiu, Bingtangcheng, Summer orange, and Navel in Clade 3. There is no evidence of clustering by geographic location suggesting that the phylogenetic lineages were mainly associated with the type of host plant regardless of sampling location³⁶.

Saturation analysis results, conducted using DAMBE v.5.3.8³⁷, indicate no significant saturation for all three codon positions ($P < 0.0001$; Iss values $<$ Iss.c values). Neither the single gene nor concatenated data sets showed any sign of pervasive homoplasy ($CI > 0.5$).

Parsimonious analysis of the concatenated mtDNA sequence yielded three equally parsimonious trees, from which a strict consensus tree was constructed. The topology of consensus tree is presented as an unrooted network with scaled branches (data not shown).

Based on the 26 identified *D. citri* mtDNA haplotype sequences, branch lengths in terms of both synonymous (syn) and non-synonymous (nonsyn) sites for each mitochondrial gene were estimated and shown in Fig. S1. The mtDNA diversification as defined by at nonsyn sites exhibited few hierarchical structures (Fig. S1). At syn sites, all genes almost showed no divergence across the entire tree. There was no evidence of phylogenetic structuring by either host plant or sampling location which is inconsistent with the NJ tree from SSR data.

Source of variation	df	SS	VC	%V	F	P
SSR analysis						
Among 3 host plants groups						
Among groups	2	417.24	0.312	10.83	$F_{CT} = 0.108$	0.001
Among populations within groups	7	206.06	0.155	4.784	$F_{SC} = 0.063$	0.004
Within populations	752	5059.14	2.413	84.41	$F_{ST} = 0.025$	0.004
Among 7 locations populations						
Among populations	6	323.809	0.213	7.83	$F_{SC} = 0.041$	0.027
within populations	1790	4418.753	0.246	92.17	$F_{ST} = 0.052$	0.034
Concatenated mtDNA analysis						
Among 3 host plants groups						
Among groups	2	10.88	0.011	3.49	$F_{CT} = 0.034$	0.101
Among populations within groups	74	33.07	0.034	41.12	$F_{SC} = 0.011$	0.055
Within populations	178	57.99	0.051	55.39	$F_{ST} = 0.005$	0.215
Among 7 locations populations						
Among populations	6	60.43	0.045	58.54	$F_{SC} = 0.045$	0.005
within populations	199	46.49	0.034	41.46	$F_{ST} = 0.034$	0.104

Table 2. AMOVA in *D. citri* populations categorized by either host plant or location and corresponding fixation indices based on SSR markers. df, Degrees of freedom; SS, sum of squares; VC, variance components; %V, percentage of variation; F, multilocus F-statistic; F_{CT} , Variation among groups; F_{SC} , Variation among populations within groups; F_{ST} , variation within population.

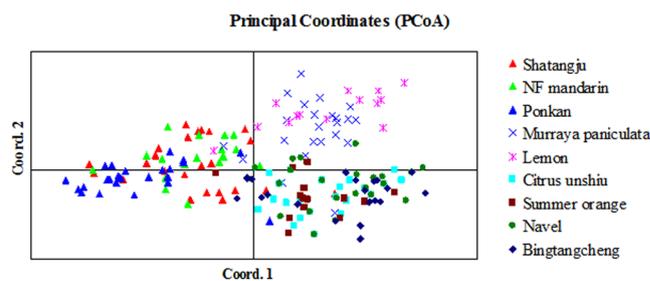


Figure 2. Principal coordinate analysis (PCoA) of pairwise distances between individuals of *D. citri* with SSR markers. Host plant populations are indicated by different colors.

Genetic diversity and population structure based on SSR data. The AMOVA analysis result revealed over 80% variance came from individuals within population which indicated that great genetic variation existing within the populations. The partitioning of total genetic variation in three host plant groups using AMOVA indicated 10.83% genetic diversity among groups, 4.78% among host plant populations (Table 2). AMOVA also revealed significant F-statistics for each hierarchical level (Table 2), suggesting that *D. citri* has strong and significant genetic differentiation at the host plant level. In contrast, AMOVA detected 7.83% of variation among locations populations and 92.17% within locations populations (Table 2), hinting very low genetic differentiation at the location level.

Principal coordinate analyses showed that the first two components jointly explained 75.6% of the genetic variation of the *D. citri* populations (right panels, Fig. 2). The first component (bottom right, Fig. 2) separated *D. citri* on Citrus unshiu, Bingtangcheng, Summer orange and Navel from the remaining, whereas the second component (top right, Fig. 2) separated *D. citri* on Murraya paniculata and Lemon from the rest of the remaining (Fig. S2). PCoA analyses based on sampling location found no clear structure or differences (Fig. S3).

The *D. citri* populations were grouped according to their host plants. Estimates were made of the genetic diversity present among host plant populations. The expected heterozygosity (H_E) in the overall *D. citri* populations' data set was 0.523, with an overall observed heterozygosity (H_O) of 0.492. Four host plant populations (Citrus unshiu, Summer Orange, Navel and Bingtangcheng) showed higher observed heterozygosity than expected; although in many cases this difference was not significant (Table 3). Mean H_E values ranged from 0.432 (Ponkan) to 0.591 (Murraya paniculata), and mean H_O ranged from 0.377 (NF mandarin) to 0.617 (Summer Orange) (Table 3). The percentage of polymorphic loci per host plant population ranged from 88.9% to 100% (Table 3).

The relationship among the nine host plant populations was further examined using the genetic distances as shown in Table 4. The overall F_{ST} among all host plant populations of *D. citri* was equal to 0.13, significantly different than zero ($p = 0.01$). Pairwise F_{ST} values ranged from 0.001 (Navel to Bingtangcheng) to 0.398 (Ponkan to Lemon) (Fig. 3A and Table 4), and comparisons of Ponkan to the other and Lemon to the other were higher than any host plant population comparison.

Population	N	N _A	H _O	H _E	Nei's	N _p	P (%)
Shatangju	32	53	0.425 ± 0.365	0.494 ± 0.294	0.486 ± 0.288	9	100
NF mandarin	22	51	0.377 ± 0.347	0.453 ± 0.277	0.441 ± 0.270	8	100
Ponkan	37	42	0.404 ± 0.371	0.432 ± 0.274	0.415 ± 0.270	8	88.89
Murraya paniculata	25	56	0.391 ± 0.272	0.591 ± 0.207	0.580 ± 0.203	9	100
Lemon	22	51	0.498 ± 0.248	0.587 ± 0.241	0.573 ± 0.235	9	100
Citrus unshiu	22	46	0.550 ± 0.323	0.540 ± 0.271	0.526 ± 0.263	9	100
Summer Orange	22	41	0.617 ± 0.408	0.510 ± 0.280	0.498 ± 0.273	8	88.89
Navel	23	41	0.578 ± 0.356	0.548 ± 0.235	0.536 ± 0.230	9	100
Bingtangcheng	20	39	0.595 ± 0.317	0.556 ± 0.218	0.542 ± 0.212	9	100

Table 3. Genetic diversity indices of *D. citri* from different host plants based on SSR markers. N, Number of samples; N_A, Observed number of alleles; Nei's, Nei's gene diversity; H_O, Observed heterozygosity; H_E, Expected heterozygosity; N_p, Number of polymorphic loci; P, Percentage of polymorphic loci;

Population	Shatangju	NF mandarin	Ponkan	Murraya paniculata	Lemon	Citrus unshiu	Summer orange	Navel	Bingtangcheng
Shatangju									
NF mandarin	0.053								
Ponkan	0.019	0.066							
Murraya paniculata	0.230	0.096	0.236						
Lemon	0.069	0.129	0.398	0.158					
Citrus unshiu	0.192	0.251	0.218	0.157	0.373				
Summer orange	0.198	0.276	0.227	0.211	0.330	0.065			
Navel	0.224	0.317	0.225	0.171	0.324	0.090	0.023		
Bingtangcheng	0.220	0.312	0.220	0.157	0.286	0.076	0.022	0.001	

Table 4. Nei's genetic distance (below diagonal) between nine host plant populations of *D. citri*.

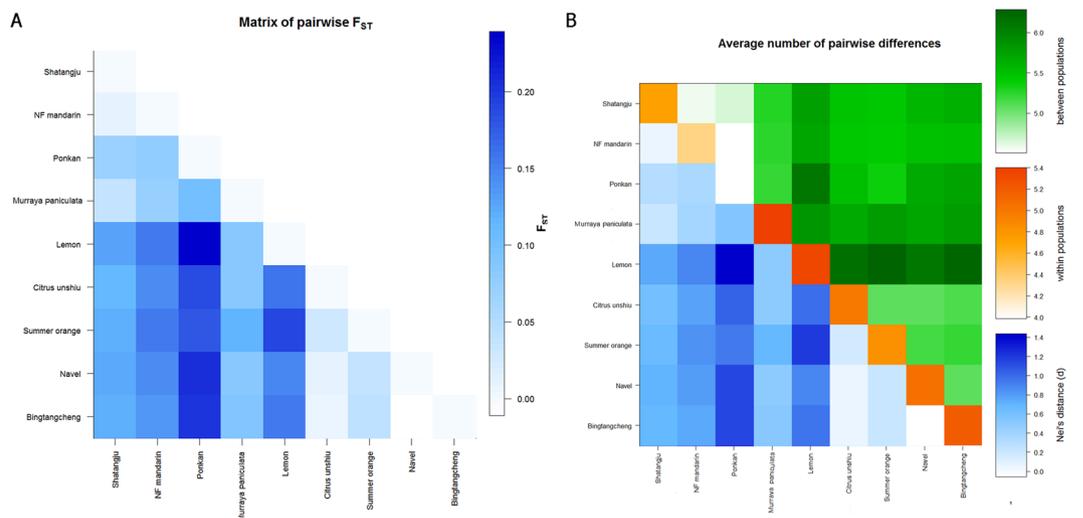


Figure 3. Graphic representation of *D. citri* host plant population relationships as described by average number of a matrix of pairwise F_{ST} values and pairwise differences (π). (A) Population relatedness represented by pairwise F_{ST} (<0.1: minimal genetic differentiation; between 0.1 and 0.2: moderate genetic differentiation; >0.2 high differentiation); (B) Population differences measured by pairwise differences π (within population on the diagonal, color coded from orange to red; between populations above the diagonal, color coded green) and genetic distance between populations (below the diagonal, blue).

We also generated estimates of the variance within vs. between host plant populations using pairwise differences (π) as a measure of population subdivision (Fig. 3B). Each of these populations showed within population heterogeneity (Fig. 3B). High pairwise differences were found between the host plant populations indicating genetic differentiation was present among *D. citri* (color coded green in Fig. 3B).

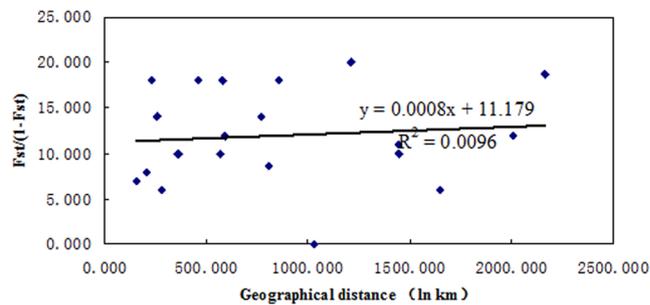


Figure 4. Scatter plots of genetic distance vs. geographical distance for pairwise population comparisons ($r^2 < 0.01$, $p \geq 0.16$; 10 000 randomizations).

Pop ID	S	Nh	Hd	π	<i>D</i>	<i>F_s</i>	<i>R</i> ²
Total Data	52	26	0.4943	0.00068	-2.61**	-5.23**	0.0139**
Host plant							
Shatangju	8	6	0.4113	0.0006	***	***	***
NF mandarin	3	4	0.6623	0.0005	***	***	***
Ponkan	22	3	0.1630	0.0012	***	***	***
Murraya paniculata	3	4	0.3477	0.0003	***	***	***
Lemon	8	5	0.6753	0.0017	***	***	***
Citrus unshiu	3	4	0.4503	0.0003	***	***	***
Summer orange	2	3	0.6381	0.0006	***	***	***
Navel	5	5	0.0003	0.0006	***	***	***
Bingtangcheng	2	2	0.1895	0.0003	***	***	***
Location							
Guangdong	1	2	0.1333	0.0001	-0.5290	0.4045	0.1777
Guangxi	2	3	0.5952	0.0005	0.03060	-0.2514	0.1544
Yunnan	9	6	0.6230	0.0014	-0.4702	0.0975	0.0781
Fujian	1	2	0.1333	0.0001	-1.7490	-1.3822	0.1588
Jiangxi	10	9	0.6113	0.0007	-1.1595	-1.5427	0.0730
Hunan	8	8	0.3778	0.0003	-2.6081**	-2.0698	0.0499
Zhejiang	24	5	0.3810	0.0012	-2.5209**	-4.4112*	0.0585*

Table 5. Summary of nucleotide polymorphism and neutrality test parameters of *D. citri* populations. S, number of segregating sites; Nh, Number of haplotypes; Hd, Haplotype diversity; *D* and *F_s*, Tajima's *D* test and Fu's *F_s*, respectively; *R*², Ramos-Onsins and Rozas' *R*² test; * $p < 0.05$; ** $p < 0.001$.

Genetic diversity and population structure based on mtDNA. Nucleotide (π) and haplotype (Hd) diversity levels calculated with syn sites was all lower than nonsyn sites as shown in Fig. S4, this is because syn sites are fast evolving and more responding to fast changing environmental factors and hence under selection. Nonsyn sites may be slow evolving and under less selective pressure. Therefore, the following analyses were all based on nonsyn sites in the concatenated mtDNA.

Resultant tree based on the concatenated mtDNA sequence did not show any significant spatial structuring due to host plant species or sampling location. The AMOVA of the concatenated mtDNA analysis revealed that 41.46% of genetic variation could be explained by the variation within populations, whereas the remaining (58.54%) came from variation among populations. Further Mantel tests of isolation by distance revealed no significant relationship between geographical and genetic distances ($r^2 < 0.01$, $p \geq 0.32$) (Fig. 4). No population structure in slow sites in mtDNA may be consistent with the fact that this species in China has only a relative short history.

The strongest association of haplotype diversity with sampling location was seen in Yunnan province (Hd: 0.62299) (Table 5 and Fig. 5A). Yunnan is a region with diverse climatic types in China and thus could be an ideal refuge for old and new *D. citri* haplotypes. The lowest haplotype diversity (Hd: 0.13333) was detected in Fujian and Guangdong provinces, which seems to contradict the fact that they are among the earliest invaded regions in China. The significantly low diversity level may be due to excessive pesticide use and removal of HLB-infected trees for pest control, which caused a higher citrus distribution disturbance and lower overall *D. citri* richness³⁸ (Table 1). Jiangxi province is a newly invaded *D. citri* area with high haplotype diversity (Hd: 0.61134), possibly hinting at multiple invasive events or that intra- and/or inter-specific hybridization exists³⁹ (Table 5). The strongest association of haplotype diversity with host species was seen in NF mandarin, Lemon and Summer orange

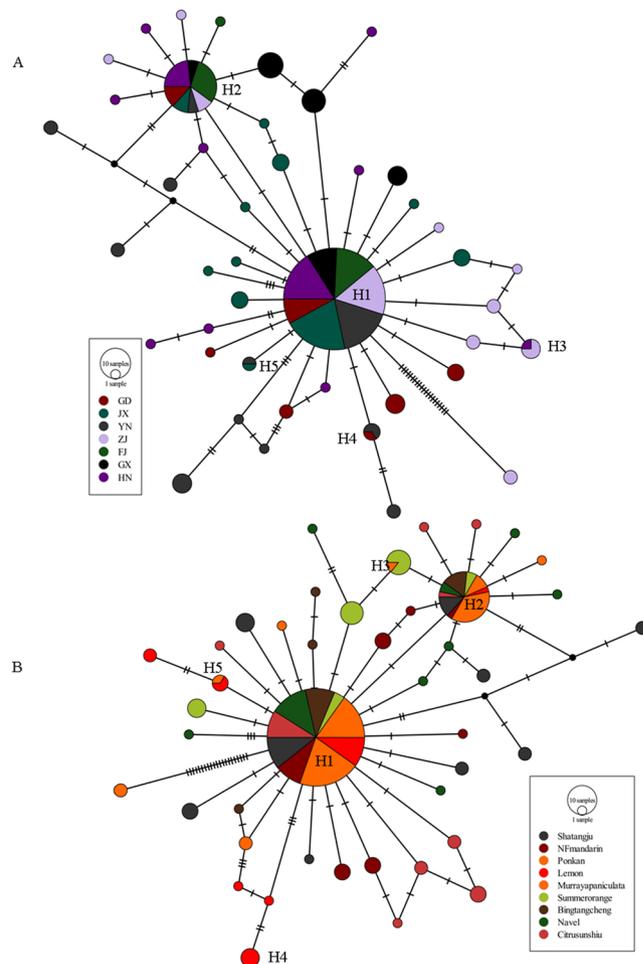


Figure 5. Median-joining networks constructed with the concatenated mitochondrial sequences, represented by (A) sampling location and (B) host plant. Circle sizes are proportional to the haplotype frequency. Distances are proportional to mutation number between haplotypes.

sampled from Jiangxi, Yunnan, and Guangxi provinces, respectively, which is consistent with the location association analysis results.

Haplotype network based on nonsyn sites. Based on the above results, we further assessed the associations of 26 unique haplotypes (Fig. 5). The overall pattern was star-like, dominated by the most common and ancient haplotypes (H1, H2), surrounded by the recent haplotypes. The star-like network suggested that most haplotypes differed from H1 and H2 by only a few mutations. The network suggested that *D. citri* experienced population expansion events. The haplotype networks, together with the haplotype distribution information (Fig. 5), indicated the ancestral haplotypes (i.e., H1 and H2) were widespread in the distribution range of *D. citri*, suggesting that a large gene pool was historically present and adequate gene exchange existed among the locations and host plants (Fig. 5).

The *D. citri* population of Yunnan (YN) had a relatively high proportion (80%) of shared haplotypes (H1, H2, H3, H4 and H5), whereas other populations (HN, GD, HN, FJ, ZJ and JX) had 60% shared haplotypes. This would suggest that the *D. citri* population from Yunnan might have established relatively early in these areas. More haplotypes (Nh) were found in Jiangxi (JX) and Hunan (HN), suggesting the existence of more relatively suitable habitats in those two regions (Table 5).

In combined analysis with host species, we observed that *Murraya paniculata* had a relatively high proportion (80%) of shared haplotypes (H1, H2, H3, H4 and H5) suggesting that *Murraya paniculata*, which is a widespread plant, is a much preferred host plant for *D. citri*. In many regions of tropical China (Yunnan, Guangxi), pruned hedges of *Murraya paniculata* are flushing all the year round, which will constitute an extreme psylla reservoir for *D. citri*⁴⁰. This would be the reason why the high haplotype frequency of *D. citri* occurred in *Murraya paniculata*.

Demographic history. Population demographic history was reconstructed using three neutrality tests and mismatch distribution analysis from nonsyn sites. Results from neutrality tests showed significant negative values for Tajima's *D* and Fu's *F_s* when all individuals were tested as a single group (Table 5). The results showed a clear deviation from neutrality (*D* and *F_s*) and were further confirmed by the Ramos-Onsins and Rozas' *R²*

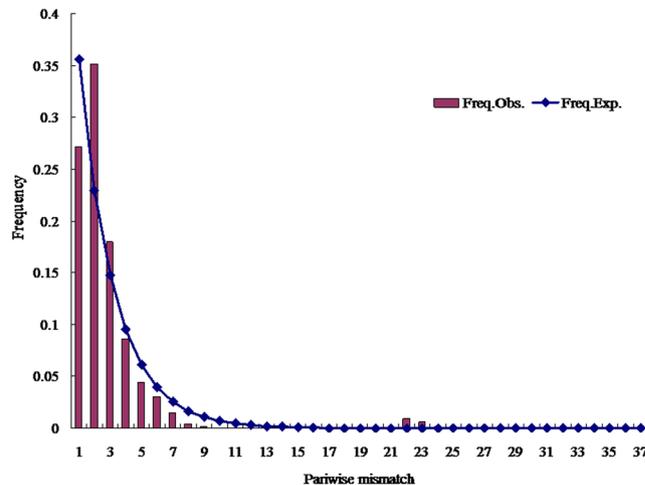


Figure 6. Distribution of pairwise nucleotide differences (mismatch distribution) of the concatenated genes in *D. citri* populations.

test (Table 5). This test gave significant results, rejecting the null hypothesis of constant size and supporting a recent demographic expansion. The unimodal mismatch distribution obtained for *D. citri* populations (Fig. 6) also suggests *D. citri* in China underwent a process demographic expansion. Statistical analysis of the mismatch distribution supports the model of demographic (SSD = 0.001; P = 0.324) expansion. Moreover, the values of *D. citri* populations size before ($\theta_0 = 0.002$) and after ($\theta_1 = 12.415$) the demographic expansion of the populations were calculated.

Also, the applied neutrality tests revealed significant deviations from neutrality for Zhejiang population (Table 5), suggesting that *D. citri* in Zhejiang province have experienced a recent expansion event.

Discussion

Overall, our present study explores nine polymorphic SSR loci and demonstrates that host-associated differentiation exists among *D. citri* populations in China. This result is derived from the SSR marker analyses, while the mtDNA sequences did not detect any significant genetic differentiation for *D. citri* populations. Furthermore, our results suggest recent population expansions occurred for *D. citri* populations. These results provide new perspectives for *D. citri* control, specifically with respect to host-associated structure among *D. citri* populations.

Molecular marker resource for *D. citri*. Even though *D. citri* is the most invasive species on citrus^{41,42}, there were only a small number of molecular markers (i.e., COI gene and 33 SSR loci) available to study its genetic diversity and population structure^{34,35}.

The SSR primer pairs designed in this study from *D. citri* genome sequence information⁴³ successfully screened nine polymorphic SSR loci, which were sufficiently variable to allow detailed studies of genetic diversity, LD and population structure. After sequential Bonferroni adjustment⁴⁴, three loci (Dci-8, Dci-29 and Dci-33) in the Chinese *D. citri* populations significantly deviated from HWE ($p < 0.05$), most likely due to heterozygote deficiency or population structure within the individuals. These newly developed markers are likely to be a valuable resource for further population genetic studies of *D. citri* from China and other countries, which could facilitate the conservation and management of *D. citri*.

Genetic structure of *D. citri* populations. There have been other studies on the genetic diversity of *D. citri* with mtDNA like COI gene. De León *et al.*¹⁸ identified 23 haplotypes from populations collected in America, while the worldwide haplotype diversity of *D. citri* reported by Boykin *et al.*¹⁹ is even lower (only 8 haplotypes). In Brazil, Guidolin *et al.*²⁰ detected 47 haplotypes using an almost complete sequence of COI gene, indicating a high haplotype diversity ($H = 0.839$) but still low in nucleotide diversity ($\pi = 0.002$). The genetic diversity and tree topology calculated further with nonsyn and syn sites in mtDNA showed a clear pattern of more variation distributed in nonsyn sites (slow evolving gene). Seemingly, neutral processes played a minor role in syn sites (fast evolving gene). Based on MGD theory, we used nonsyn sites for *D. citri* phylogenetic analysis. Genealogy indicated a limited intraspecific variation in *D. citri* population from China showing no clear grouping cluster. Different from mtDNA, the results from the SSR markers for the same *D. citri* populations showed the presence of three host-associated genetically distinct clusters from tree topology and PCoA analysis, which were not related to geographic distance among the *D. citri* populations. Those SSR data suggested that the populations of *D. citri* may be experiencing some degree of host plant-associated differentiation. Nevertheless, the phylogenetic trees constructed from the SSR data could be biased towards host plant selection because the SSR markers were derived mostly from fast evolving sites. Cautions are therefore recommended when interpreting the phylogenetic analysis results⁴⁵.

Such discrepancies between different types of markers are also reported in *Procambarus Clarkii*, *Syngnathus typhle* and *Clupea harengus* population genetic studies^{46–48}. The inconsistent results obtained from the two

molecular markers are generally explained by the features of the two kinds of markers. SSRs, with their high mutation rate of 10^{-2} to 10^{-6} /locus/generation^{49,50}, can increase the overall resolution to detect genetic differentiation and give clues about recent and contemporary events in the population⁵¹. By comparison, the slower mutation rate of the mtDNA sequence marker may provide more of a historic than a recent picture of gene flow⁵². On the other hand, such results are consistent with the Maximum Genetic Diversity (MGD) hypothesis^{22,36,53}. SSRs as the fast evolving sequences are at saturation genetic diversity (rather than neutral) that is maintained by stabilizing natural selection^{36,53}. In our case, host plants provide the selection force. Sequences not compatible with living in a particular plant species would be selected out. Slow evolving sites (mtDNA) are too slow to meet adaptive needs and hence appear neutral within short time frames⁵³.

Host-associated selection driving population genetic differentiation. In this study, *D. citri* populations were found to be structured into three subgroups by host plant: (i) Shatangju, NF mandarin and Ponkan; (ii) *Murraya paniculata*, Lemon; (iii) Citrus unshiu, Bingtangcheng, Summer orange, and Navel. Mechanisms that may contribute to host-associated differentiation of *D. citri* populations were explored based on the taxonomic, volatile and phenological differences among hosts.

Numerous classification systems of Citrus have been formulated, among which that of Swingle⁵⁴, who described 16 species in Citrus, has been the most widely accepted⁵⁵. Shatangju, NF mandarin, Ponkan and Citrus unshiu all belong to *Citrus reticulata*, while Lemon belongs to *Citrus medica*; Bingtangcheng and Summer orange belong to *Citrus sinensis*. We noticed that the Citrus classification is in accordance with the genetic structure of *D. citri* populations besides Citrus unshiu. Plants are the site of assembly for *D. citri*. Population fluctuations of *D. citri* on citrus and other plants are closely correlated with the rhythm, quantity, and nutritional quality of plant flush because eggs and nymphs are exclusively associated with new growth⁵⁶. Vassilou^{57,58} found that Lemon (*Citrus medica*) trees flush sporadically throughout the year, providing an important feeding source and refuge for *D. citri* over winter. *Murraya paniculata* has a more cold-hardy bud break, advancing flushing by six weeks⁵⁸. When compared to *Citrus sinensis*, *Citrus reticulata* trees are more winter-hardy, letting the autumn plant flush extend to November⁵⁹. The prolonged flushing stage would extend the lifespan of *D. citri*. Furthermore, whether or not host plants are available at the early spring and the whole winter poses another high selection pressure on insect populations through effects on diapause^{60,61}. Variation in diapause would contribute to the divergence of host-associated populations of insects^{62,63}, as reported in *Rhagoletis pomonella*⁶⁴ and *Diatraea saccharalis*⁶¹. The flushing characteristics of Lemon and *Murraya paniculata* described above could make diapause termination be more beneficial to their *D. citri* populations than populations on other citrus species. Therefore, we suggest that host plant phenological differences may contribute to host-associated differentiation in *D. citri* populations.

Host plants provide insects not only nutrients but also secondary substances such as volatiles influencing insect behavior. Chemical and nutritional differences existing among host plants may be a reason for *D. citri* host-associated differentiation. Ehrlich & Raven⁶⁵ found chemical similarities among unrelated food plants used by related butterflies, and suggested that secondary plant substances played the leading role in determining patterns of utilization. Aroma compounds are important for citrus; each cultivar type generates a different profile of volatiles that contributes to its distinct aroma attributes⁶⁶. These volatiles provide a critical clue for host plant selection and orientation in *D. citri*. Linalool was reported to be a major constituent among leaf volatiles in young leaves of most cultivars⁶⁷. γ -terpinene was a major component in leaf oil of *Citrus sinensis*^{68,69} and thymol ethyl ether was only found in leaf volatiles of *Citrus reticulata*^{70,71}. Limonene was predominantly found in *Citrus medica* (lemon), and β -bisabolene was only identified in *Citrus medica* leaves⁷². Interestingly, β -bisabolene was also detected in leaf volatiles of *Murraya paniculata* oil⁷³. Therefore, such chemical differences among *D. citri* host plant species could act as an internal driver of locally adapted populations as reported by Ehrlich⁶⁵.

Insects can develop an oviposition preference for the natal plant^{63,74}, and assortative mating could occur between individuals on the same host plant⁷⁵. Several lines of evidence have suggested that female host preferences exist in *D. citri*⁷⁶. Investigation of the location and preference of adult *D. citri* on *Murraya paniculata*, Lemon (*Citrus medica*) and Ponkan (*Citrus reticulata*) found no difference between adult numbers on *Murraya paniculata* and Lemon, but significantly fewer on Ponkan, suggesting that *Murraya paniculata* and Lemon are the preferred hosts⁷⁷. Field observations showed that *Citrus medica* suffered the heaviest from *D. citri*, followed by *Citrus sinensis*, with *Citrus reticulata* the lightest^{78,79}. Female *D. citri* orient to their plant of origin; males orient to the odor of females and co-specific excretions⁸⁰. This behavior in combination with female oviposition preference behavior likely results in mate choice and limits gene flow between *D. citri* from different host plant clusters^{81,82}. In addition, micro-ecological variables such as symbiont composition among different host plants of *D. citri* may also be an important source of host plant-associated natural selection. Study of the population structure and symbiont community of host-adapted *Acyrtosiphon pisum* populations showed that the presence of six facultative (secondary) bacterial endosymbionts was nonrandom distributed across aphid genetic clusters and suggested that bacterial endosymbionts affect population divergence and speciation⁸³. More experiments are needed to uncover whether micro-ecological variables exist in host plant-associated *D. citri* populations.

Demographic history of populations. Both neutrality tests and unimodal mismatch distributions revealed an expansion of *D. citri* populations. The recent population expansion of *D. citri* and its increasing damage around the world may have resulted from the recent sharp expansion of citrus varieties and growing area, which has provided more food options and ecological environments for *D. citri*⁸⁴. Beck & Reese⁸⁵ proposed that insect survivorship, fecundity, growth rate and activity can be affected by the quantity and quality of hosts. An abundance of hosts in a habitat will increase survival and fecundity and reduce mortality. As *Murraya paniculata* is the primary wild host for *D. citri*, the increase in population size of this species would distinctly improve the fecundity and survival capacity of the insect^{13,86}, and the extensive commercial cultivation of citrus trees in recent years may also have contributed to an expansion of the insect's territory⁸⁴. Additionally, the warmer temperatures

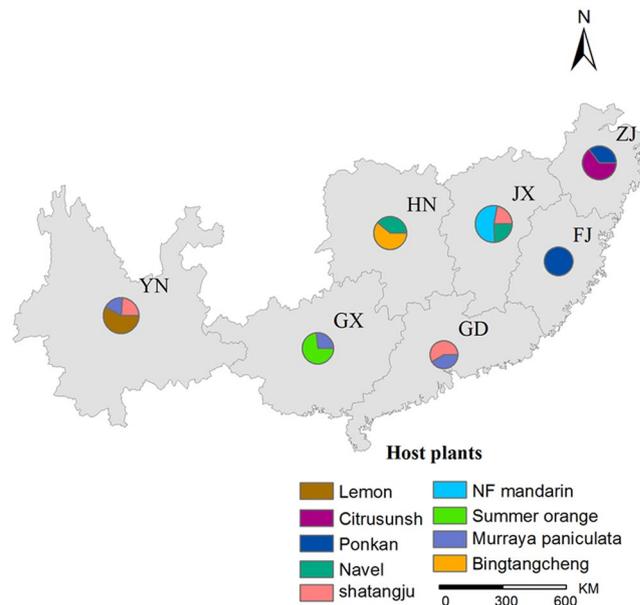


Figure 7. Geographical sites and host plants of sampled *D. citri* in China. The map was drawn with Arcgis 10.2 software (<http://www.esri.com/arcgis/about-arcgis>).

Province	Location	Number	Geographical coordinates	Host plant
Zhejiang (ZJ)	Lishui	12	28°26'N 119°54'E	Ponkan
	Taizhou	22	28°38'N 121°09'E	Citrus unshiu
Hunan (HN)	Chenzhou	7	25°22'N 112°53'E	Navel
	Yongzhou	6	25°37'N 111°31'E	Navel
	Chenzhou	20	26°30'N 113°30'E	Bingtangcheng
Jiangxi (JX)	Fuzhou	22	27°12'N 116°30'E	NF mandarin
	Ganzhou	10	25°50'N 114°55'E	Navel
	Ganzhou	9	25°17'N 114°57'E	Shatangju
Fujian (FJ)	Quanzhou	25	25°23'N 118°16'E	Ponkan
Yunnan (YN)	Baoshan	22	24°58'N 98°52'E	Lemon
	Mangshi	9	24°26'N 98°26'E	Shatangju
	Ruili	7	24°3'N 97°56'E	Murraya paniculata
Guangxi (GX)	Guilin	4	25°16'N 110°19'E	Murraya paniculata
	Liuzhou	4	24°39'N 109°14'E	Murraya paniculata
	Guilin	22	24°24'N 110°23'E	Summer orange
Guangdong (GD)	Zhaoqing	10	23°8'N 112°5'E	Murraya paniculata
	Zhaoqing	7	23°2'N 112°23'E	Shatangju
	Guangzhou	7	23°9'N 113°20'E	Shatangju
Total		225		9

Table 6. Details of *D. citri* samples used in the study.

and increased precipitation not only contribute to *D. citri*'s physiological activity but also provide suitable living conditions for *D. citri* in relatively dry seasons in South China, increasing the number of habitats in Mainland China^{6,87}.

Materials and Methods

Sample collection. A total of 225 adult *D. citri* individuals were collected from 9 host plant types located in 20 citrus production regions across 7 governmental provinces in China (Fig. 7 and Table 6). Upon sampling, geographical coordinates and host plant species were recorded. To maximize the coverage of the genetic variation among samples, we collected *D. citri* from an additional 10 host plants within a 10-mile zone centered on each citrus sampling site per region. Samples were stored in 70% ethanol at 4 °C until DNA extraction.

DNA extraction and PCR amplification. Total genomic DNA was extracted from a single intact insect following a modified salt-extraction protocol⁸⁸. The DNA was eluted in 40 mL TE buffer and stored at -20°C . Final concentrations and 280/260, 260/230 ratios were estimated with a NanoDrop UV spectrophotometer (Techno Scientific, Wilmington, DE, USA). DNA amount per sample was normalized to 30 ng/ μL . The polymerase chain reaction (PCR) protocols for SSR markers and mtDNA genes were as follows:

SSR amplification and genotyping. Based on the entire genome, SSR Hunter 1.3 software⁸⁹ was employed for SSR content. Primer pairs were designed in Software PRIMER v 3.0⁹⁰ with sufficient flanking sequence and tested for amplification and polymorphism using DNA obtained from thirty individuals. The set of individuals was the same as those used for variability testing (see below) and individuals were rotated during initial screening such that a different set of thirty was used for each primer pair. The variability of the selected polymorphic loci was assessed in 225 specimens, and they were deposited in GenBank with the Accession nos. KY053427-KY053435. The conditions and characteristics of the SSR loci are provided in Table 1. For SSR genotyping, all forward primers were tagged with M13 (-21) ($5'$ -TGAAAACGACGGCCAGT- $3'$) at the $5'$ end. PCR amplifications were performed in a 10 μL reaction volume (4.5 μL Takara Master Mix [Shanghai Corporation], 0.4 μM of both forward and reverse primers, 0.2 μM dye labeled primer, 0.1% Triton X-100, 1.2 μL water and 1 μL DNA template [$\sim 30\text{ng}$]). The PCR conditions were: initial denaturing at 95°C for 2 min, followed by 35 cycles at 95°C for 30 s, annealing at primer specific temperature for 45 s, extension at 72°C for 45 s. Then, the reaction was paused at 72°C , and fluorescently labeled (6-FAM, HEX or TAMRA) M13 primer was added. Amplification was continued for the next 10 cycles according to De Arruda⁹¹ with 30 min of final extension at 72°C . Fragment analyses were run on an ABI 3730 DNA sequencer (Applied Biosystems) and analyzed with GENEMARKER 2.2.0 (Soft Genetics) using GS-500 (LIZ) as a size standard.

Sequencing of SSR loci was performed for the presence of size homoplasy detection. To isolate the specific bands of SSRs, the PCR products were separated on a 2.5% agarose gel. Only the best amplified bands for each SSR primer combination was collected and purified with QIA quick PCR Purification Kit (Omega). Those DNA fragments were sequenced with the specific forward or reverse SSR primer on an ABI3700 DNA analyser. The sequences obtained were checked for their homology with the SSR source sequence.

Mitochondrial DNA amplification. Three mitochondrial genes (cytochrome c oxidase subunit 1, COI (781 bp); cytochrome b, Cytb (867 bp); and NADH dehydrogenase subunit 5, ND5 (750 bp)) were amplified and sequenced. Products were amplified by PCR in 25 μL reactions with 1.25 units of AmpliTaq DNA polymerase, 1 \times PCR buffer, 2 mM of MgCl_2 , 0.4 mM of dNTPs and 0.5–1 μM of each primer. A detailed description of the primers used and PCR conditions can be found in the Supporting Information (Table S1). Sequencing was carried out using fluorescently labeled dideoxy terminators on an ABI 3130xl Genetic Analyzer (Applied Biosystems Corp.).

Data analyses. *SSR analysis, genetic diversity.* We employed the Quantity One software (Biorad, USA) to estimate the band sizes adjusted for the standard molecular weight size marker (20 bp DNA ladder, TAKARA) and then formatted these size-based data to be used in GenAlEx v 6.0⁹² for further analysis. The genotypic data were used to estimate genetic diversity. Allele dropout and null alleles in the newly developed SSR data were checked using MICROCHECKER v 2.2.3⁹³. The observed number of alleles (N_A), observed heterozygosity (H_o), expected heterozygosity (H_e), percentage of polymorphic loci (P), Nei's genetic diversity (Nei's), and Nei's original measure genetic distance (Nei's D) were calculated with GenAlEx v 6.0 software. Deviations from Hardy-Weinberg equilibrium (HWE) and linkage equilibrium between loci were quantified in GENEPOP v 1.2⁹⁴ within populations. The inbreeding coefficient (F_{IS}) and confidence intervals were calculated in FSTAT v 1.2 with 1000 random permutations⁹⁵.

Phylogenetic genealogy construction with SSRs. Unrooted neighbor-joining trees were constructed using matrices of chord genetic distance⁹⁶ using the PHYLIP v3.5c⁹⁷ software. The CONSENSE subroutine within PHYLIP was used to perform 1000 bootstraps of the distance matrix. The resulting trees were visualized using TREEVIEW v 1.6.6⁹⁸.

Population structure based on SSRs. Different analyses were conducted to examine the contemporary population structure of *D. citri* populations in China. First, principal coordinate analysis (PCoA) was performed based on the simple matching coefficient using the decenter and eigenvector matrices in GenAlEx v 6.0 with 1000 random permutations. Second, analysis of molecular variance (AMOVA) was performed, where the total genetic variability was subdivided into three variance components, i.e., among clusters (F_{CT}), among populations within clusters (F_{SC}), within populations (F_{ST}) using the software Arlequin v 3.5 with 1000 permutations^{99,100}.

Genetic diversity and population structure with mtDNA. The 225 mtDNA sequences were compiled initially in Clustal-X¹⁰¹ and further refined with BioEdit 7.0¹⁰². Haplotype diversity (H_d) and nucleotide diversity (π) of total *D. citri* individuals were calculated as defined by Nei¹⁰³ in DnaSP v 5.0¹⁰⁴ with nonsyn and syn sites separately. Genetic parameters for each population grouped by either location or host, including number of haplotypes (N_h), haplotype diversity (H_d) and nucleotide diversity (π), and number of segregating sites were further analyzed with nonsyn sites in concatenated mtDNA. AMOVA analysis was performed in the same way as described for the SSR data analysis above to test for whether significant population structure exists using the HKY model¹⁰⁵. Evidence of isolation by distance among sampling locations in China was obtained by examining the correlations between matrices of pairwise genetic distance (F_{ST}) and pairwise physical distance¹⁰⁶. Mantel tests¹⁰⁷ were used to test the strength of the correlations between the matrices in GenAlEx v 6.0.

Haplotype phylogeny with mtDNA. Substitution saturation was assessed for the alignments of single genes and concatenated dataset with DAMBE v. 5.3.46³⁷. To measure the level of homoplasy, the consistency index (CI) was calculated using PAUP* v4.0b10¹⁰⁸. Phylogenetic analysis with the maximum parsimony (MP) method was carried out by PAUP* v4.0b10 using 100 replicates, MAXTREES = 200,000, random addition, TBR branch swapping, and MULTREES, with gaps treated as missing data. Support for the monophyletic groups revealed in the maximally parsimonious trees was examined using 1000 bootstrap replicates for the bootstrap (BP) consensus tree. In terms of syn and nonsyn substitutions per site of mtDNA, branch lengths of each mitochondrial gene were estimated with 26 previously defined sequences using a codon-based model of substitution within the codeml application in PAML v4.0¹⁰⁹. Tree topologies were visualized using TREEVIEW v 1.6.6⁹⁸. Further, we constructed a haplotype network with nonsyn sites in the concatenated mtDNA to visualize relationships among haplotypes in the program POPART based on a median-joining algorithm¹¹⁰ designed for populations of closely related samples allowing persistent ancestral nodes and multi-furcation in the network as extra information for population genetic structures.

Demographic history *D. citri* population in China. Signatures of *D. citri* populations' demographic changes were investigated using three neutrality tests: Tajima's D ¹¹¹, Fu's F_s ¹¹², and Ramos-Onsins & Rozas' R^2 ¹¹³. All methods are premised on neutral selection. The neutral theory states that the majority of the genome is variable and neutral. Actually, this assumption is not true²². According to Huang *et al.*³⁶, the value of the theory is its description for the linear phase, which has been retained by the MGD. Hence, the following neutrality tests of *D. citri* were from the nonsyn sites. The examination of deviation from neutrality by both D and F_s indices was based on 1000 coalescent simulations, as implemented in Arlequin v 3.5. Significant negative D and F_s values can be interpreted as signatures of population expansion. The R^2 statistic of Ramos-Onsins & Rozas, which has more statistical power when population sizes are small, was calculated using a coalescent simulation algorithm implemented in DnaSP, with 1000 simulations.

To provide other estimates of population size changes, we also examined site mismatch distributions. Contrasting plots of observed and theoretical distributions of site differences yield insight into past population demographics. The expected mismatch distributions under a sudden expansion model were computed in Arlequin v 3.5. The sum of squared deviations (SSD) between observed and expected distributions was used as a measure of fit, and the probability (P) of obtaining a simulated SSD greater than or equal to the expected SSD was computed by randomization. If this probability (P) was >0.05, the expansion model was accepted. Graphical representation was implemented in Arlequin.

Data Availability. GenBank accessions of nine polymorphic loci are KY053427-KY053435.

References

- Boina, D. R., Youn, Y., Folimonova, S. & Stelinski, L. L. Effects of pymetrozine, an anti feed ant of Hemiptera, on Asian citrus psyllid, *Diaphorina citri*, feeding behavior, survival and transmission of *Candidatus*, *Liberibacter asiaticus*. *Pest. Manag. Sci.* **67**, 146–155 (2011).
- Gottwald, T. R. Current epidemiological understanding of citrus huanglongbing. *Annu. Rev. Phytopathol.* **48**, 119–139 (2010).
- Belasque, J. *et al.* Lessons from huanglongbing management in São Paulo State, Brazil. *J. Plant. Pathol.* **92**, 285–302 (2010).
- Ledford, H. Geneticists enlist engineered virus and CRISPR to battle citrus disease. *Nature* **545**, 277–278 (2017).
- Kiefl, J. *et al.* Investigation on Key Molecules of Huanglongbing (HLB)-Induced Orange Juice Off-flavor. *Agr. Food. Chem.* (2017).
- Grafton-Cardwell, E. E., Stelinski, L. L. & Stansly, P. A. Biology and management of Asian citrus psyllid, vector of the huanglongbing pathogens. *Annu. Rev. Entomol.* **58**, 413–432 (2013).
- Kuwayama, S. Die psylliden Japans. *Transactions of the Sopporo Natural History Society* **2**, 149–189 (1908).
- Lima, C. Insetos do Brasil. Escola Nacional de Agronomia (1950).
- Halbert, S. E. Entomology section. *Tri-ology Technical Report* **37**, 6–7 (1998).
- Halbert, S. E. & Manjunath, K. L. Asian citrus psyllids (Sternorrhyncha: Psyllidae) and greening disease of citrus: a literature review and assessment of risk in Florida. *Fla. Entomol.* **87**, 330–353 (2004).
- Qureshi, J. A., Rogers, M. E., Hall, D. G. & Stansly, P. A. Incidence of invasive *Diaphorina citri* (Hemiptera:Psyllidae) and its introduced parasitoid *Tamarixia radiata* (Hymenoptera: Eulophidae) in Florida citrus. *J. Econ. Entomol.* **102**, 247–256 (2009).
- Luis, M. *et al.* Occurrence of citrus huanglongbing in Cuba and association of the disease with *Candidatus* *Liberibacter asiaticus*. *J. Plant. Pathol.* 709–712 (2009).
- Yang, Y. *et al.* Distribution, biology, ecology and control of the psyllid *Diaphorina citri* Kuwayama, a major pest of citrus: a status report for China. *Int. J. Pest. Manage.* **52**, 343–352 (2006).
- Shimwela, M. M. *et al.* First occurrence of *Diaphorina citri* in East Africa, characterization of the *Ca. Liberibacter* species causing huanglongbing (HLB) in Tanzania, and potential further spread of *D. citri* and HLB in Africa and Europe. *Eur. J. Plant. Pathol.* **142**, 1–20 (2016).
- Narouei-Khandan, H. A., Halbert, S. E., Worner, S. P. & van Bruggen, A. H. Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling approaches, with emphasis on the USA. *Eur. J. Plant. Pathol.* **144**, 655–670 (2016).
- Guillemaud, T., Ciosi, M., Lombaert, E. & Estoup, A. Biological invasions in agricultural settings: insights from evolutionary biology and population genetics. *C. R. Biol.* **334**, 237–246 (2011).
- Behura, S. K. Molecular marker systems in insects: current trends and future avenues. *Mol. Ecol.* **15**, 3087–3113 (2006).
- De León, J. H. *et al.* Two separate introductions of Asian citrus psyllid populations found in the American continents. *Ann. Entomol. Soc. Am.* **104**, 1392–1398 (2011).
- Boykin, L. M. *et al.* Overview of worldwide diversity of *Diaphorina citri* Kuwayama mitochondrial cytochrome oxidase 1 haplotypes: two Old World lineages and a New World invasion. *B. Entomol. Res.* **102**, 573–582 (2012).
- Guidolin, A. S., Fresia, P. & Cónsoli, F. L. The genetic structure of an invasive pest, the Asian citrus psyllid *Diaphorina citri* (Hemiptera: Liviidae). *PloS One* **9**, e115749 (2014).
- Avise, J. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press (2000).
- Huang, S. Primate phylogeny: molecular evidence for a pongid clade excluding humans and a prosimian clade containing tarsiers. *Sci China Life Sci* **55**, 709–725 (2012).
- Timm, J., Planes, S. & Kochzius, M. High similarity of genetic population structure in the false clown anemonefish (*Amphiprion ocellaris*) found in microsatellite and mitochondrial control region analysis. *Conserv. Genet.* **13**, 693–706 (2012).

24. Csilléry, K., Blum, M. G., Gaggiotti, O. E. & François, O. Approximate bayesian computation (abc) in practice. *Trends Ecol Evol.* **25**(7), 410–418 (2010).
25. Fontaine, M. C. *et al.* Postglacial climate changes and rise of three ecotypes of harbour porpoises, *Phocoena phocoena*, in western Palearctic waters. *Mol. Ecol.* **23**, 3306–3321 (2014).
26. Perera, O. P., Snodgrass, G. L., Scheffler, B. E., Gore, J. & Abel, C. A. Characterization of eight polymorphic microsatellite markers in the tarnished plant bug, *Lygus lineolaris* (Palisot de Beauvois). *Mol. Ecol. Notes.* **7**, 987–989 (2007).
27. Woodard, S. H. *et al.* Molecular tools and bumble bees: revealing hidden details of ecology and evolution in a model system. *Mol. Ecol.* **24**, 2916–2936 (2015).
28. Vukosavljev, M. *et al.* Efficient development of highly polymorphic microsatellite markers based on polymorphic repeats in transcriptome sequences of multiple individuals. *Mol. Ecol. Resour.* **15**, 17–27 (2015).
29. Zhang, Y., He, J., Zhao, P. X., Bouton, J. H. & Monteros, M. J. Genome-wide identification of microsatellites in white clover (*trifolium repens*.) using FIASCO and phpSSRminer. *Plant Methods* **4**, 19 (2008).
30. Nybom, H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* **13**, 1143–1155 (2004).
31. Bouck, A. M. Y. & Vision, T. The molecular ecologist's guide to expressed sequence ta. *gs. Mol. Ecol.* **16**, 907–924 (2007).
32. Singh, R. & Sheoran, S. Expressed Sequence Tags (ESTs): Resource for in silico Prediction of SSR and SNP. *Applied Computational Biology and Statistics in Biotechnology and Bioinformatics* (2012).
33. Ebert, D. & Peakall, R. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Mol. Ecol. Resour.* **9**, 673–690 (2010).
34. Boykin, L. M. *et al.* Twelve polymorphic microsatellite loci from the Asian citrus psyllid, *Diaphorina citri* Kuwayama, the vector for citrus greening disease, huanglongbing. *Mol. Ecol. Notes.* **7**, 1202–1204 (2007).
35. Tomimura, K., Inoue, H. & Iwanami, T. Development of simple sequence repeat markers for the Japanese population of the Asian citrus psyllid, *Diaphorina citri* (Hemiptera: Psyllidae). *Appl. Entomol. Zool.* **49**, 353–357 (2014).
36. Huang, S. New thoughts on an old riddle: what determines genetic diversity within and between species? *Genomics* **108**, 3–10 (2016).
37. Xia, X. & Xie, Z. DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.* **92**, 371–373 (2001).
38. Fahrig, L. & Jonsen, I. Effect of habitat patch characteristics on abundance and diversity of insects in an agricultural landscape. *Ecosystems* **1**, 197–205 (1998).
39. Haliem, E. A. & Al-Huqail, A. A. Comparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis and restricted fragment length polymorphism among fenugreek accessions. *Genet. Mol. Res.* **12**, 6284–6298 (2013).
40. Aubert, B. Integrated activities for the control of huanglongbing-greening and its vector *Diaphorina citri* Kuwayama in Asia. *Proceedings of the 4th International Asia Pacific Conference on Citrus Rehabilitation* **410**, 133–144 (1990).
41. Michaud, J. P. Numerical response of *Olla v-nigrum* (Coleoptera: Coccinellidae) to infestations of Asian citrus psyllid, (Hemiptera: Psyllidae) in Florida. *Fla. Entomol.* **84**, 608–612 (2001).
42. Sache, I., Roy, A. S., Suffert, F. & Desprez-Loustau, M. L. Invasive plant pathogens in Europe. *Biological invasions: Economic and environmental costs of alien plant, animal, and microbe species*, 227–242 (2011).
43. Vyas, M. *et al.* Asian citrus psyllid expression profiles suggest *Candidatus Liberibacter asiaticus*-mediated alteration of adult nutrition and metabolism, and of nymphal development and immunity. *PLoS One* **10**, e0130328 (2015).
44. Rice, W. R. Analyzing tables of statistical tests. *Evolution* **43**, 223–225 (1989).
45. Powell, W., Machray, G. C. & Provan, J. Polymorphism revealed by simple sequence repeats. *Trends in plant science* **1**, 215–222 (1996).
46. Li, Y. *et al.* Population genetic structure and post-establishment dispersal patterns of the red swamp crayfish *procambarus clarkii* in china. *Plos One* **7**, e40652 (2012).
47. Wilson, A. B. & Eigenmann Veraguth, I. R. I. S. The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Mol. Ecol.* **19**, 4535–4553 (2010).
48. Limborg, M. T. *et al.* Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Mol. Ecol.* **21**, 3686–3703 (2012).
49. Vázquez, J. E., Pérez, T., Albornoz, J. & Dominguez, A. Estimation of microsatellite mutation rates in *Drosophila melanogaster*. *Genet. Res.* **76**, 323–326 (2000).
50. Waples, R. S. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* **89**, 438–450 (1998).
51. Borrell, Y. J., Piñera, J. A., Sánchez Prado, J. A. & Blanco, G. Mitochondrial DNA and microsatellite genetic differentiation in the European anchovy *Engraulis encrasicolus* L. *ICES. J. Mar. Sci.* **69**, 1357–1371 (2012).
52. Fauvelot, C., Bernardi, G. & Planes, S. Reductions in the mitochondrial DNA diversity of coral reef fish provides evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution* **57**, 1571–1583 (2003).
53. Yuan, D. *et al.* Modern human origins: multiregional evolution of autosomes and East Asia origin of Y and mtDNA. *bioRxiv.* **101410** (2017).
54. Swingle, W. T. The botany of citrus and its wild relatives of the orange subfamily (Family Rutaceae, subfamily Aurantioideae). *In H. J. Webber and L. D. Batchelor, the Citrus Industry* **1**, 129–474 (1943).
55. Nicolosi, E. *et al.* Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor. Appl. Genet.* **100**, 1155–1166 (2000).
56. Hall, D. G., Hentz, M. G. & Adair, R. C. Jr. Population ecology and phenology of *Diaphorina citri* (Hemiptera: Psyllidae) in two Florida citrus groves. *Environ. Entomol.* **37**, 914–924 (2008).
57. Vassiliou, V. A. Chemical control of *Pezothrips kellyanus* (Thysanoptera: Thripidae) in citrus plantations in Cyprus. *Crop. Prot.* **26**, 1579–1584 (2007).
58. Vassiliou, V. A. Ecology and behavior of *Pezothrips kellyanus* (Thysanoptera: Thripidae) on citrus. *J. Econ. Entomol.* **103**, 47–53 (2010).
59. Guo, W., Cheng, Y. & Deng, X. Regeneration and molecular characterization of intergeneric somatic hybrids between *Citrus reticulata* and *Poncirus trifoliata*. *Plant. Cell. Rep.* **20**, 829–834 (2002).
60. Danks, H. V. Insect life-cycle polymorphism: theory, evolution and ecological consequences for seasonality and diapause control (Vol. 52). *Springer Science & Business Media* (2013).
61. Joyce, A. L. *et al.* Host-plant associated genetic divergence of two *Diatraea* spp. (Lepidoptera: Crambidae) stemborers on novel crop plants. *Ecol. Evol.* **6**, 8632–8644 (2016).
62. Dickey, A. M. & Medina, R. F. Testing host-associated differentiation in a quasi-endophage and a parthenogen on native trees. *J. Evol. Biol.* **23**, 945–956 (2010).
63. Wood, T. K., Tilmon, K. J., Shantz, A. B., Harris, C. K. & Pesek, J. The role of host-plant fidelity in initiating insect race formation. *Evol. Ecol. Res.* **1**, 317–332 (1999).
64. Feder, J. L. & Filchak, K. E. It is about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. *In Proceedings of the 10th International Symposium on Insect-Plant Relationships. Springer Netherlands* 211–225 (1999).
65. Ehrlich, P. R. & Raven, P. H. Butterflies and plants: a study in coevolution. *Evolution* **18**, 586–608 (1964).

66. Azam, M., Jiang, Q., Zhang, B., Xu, C. & Chen, K. Citrus leaf volatiles as affected by developmental stage and genetic type. *Int. J. Mol. Sci.* **14**, 17744–17766 (2013).
67. Lin, S. Y., Roan, S. F., Lee, C. L. & Chen, I. Z. Volatile organic components of fresh leaves as indicators of indigenous and cultivated citrus species in Taiwan. *Biosci. Biotechnol. Biochem.* **74**, 806–811 (2010).
68. Steuer, B., Schulz, H. & Läger, E. Classification and analysis of citrus oils by NIR spectroscopy. *Food. Chem.* **72**, 113–117 (2001).
69. Njoroge, S. M., Koaze, H., Karanja, P. N. & Sawamura, M. Essential oil constituents of three varieties of Kenyan sweet oranges (*Citrus sinensis*). *Flavour. Frag. J.* **20**, 80–85 (2005).
70. Casado, D., Gemenó, C., Avilla, J. & Riba, M. Diurnal variation of walnut tree volatiles and electrophysiological responses in *Cydia pomonella* (Lepidoptera: Tortricidae). *Pest. Manag. Sci.* **64**, 736–747 (2008).
71. Njoroge, S. M., Mungai, H. N., Koaze, H., Phi, N. T. L. & Sawamura, M. Volatile constituents of mandarin (*Citrus reticulata* Blanco) peel oil from Burundi. *J. Essent. Oil Res.* **18**, 659–662 (2006).
72. Bhuiyan, M. N. I., Begum, J., Sardar, P. K. & Rahman, M. S. Constituents of peel and leaf essential oils of *Citrus medica* L. *J. Sci. Res.* **1**, 387–392 (2009).
73. Silva, J. A. *et al.* Repellency of selected *Psidium guajava* cultivars to the Asian citrus psyllid, *Diaphorina citri*. *Crop. Prot.* **84**, 14–20 (2016).
74. Diehl, S. R. & Bush, G. L. An evolutionary and applied perspective to insect biotypes. *Annu. Rev. Entomol.* **29**, 471–504 (1984).
75. Havens, J. A. & Etges, W. J. Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. IX. Host plant and population specific epicuticular hydrocarbon expression influences mate choice and sexual selection. *J. Evol. Biol.* **26**, 562–576 (2013).
76. Alves, G. R., Diniz, A. J. F. & Parra, J. R. P. Biology of the huanglongbing vector *Diaphorina citri* (Hemiptera: Liviidae) on different host plants. *J. Econ. Entomol.* **107**, 691–696 (2014).
77. Yasuda, K., Kawamura, F. & Oishi, T. Location and preference of adult Asian citrus psyllid, *Diaphorina citri* (Homoptera: Psyllidae) on Chinese box orange jasmine, *Murraya exotica* L. and flat lemon, *Citrus depressa*. *Jpn. J. Appl. Entomol. Z.* **49**, 146–149 (2005).
78. Hall, D. G., Richardson, M. L., Ammar, E. D. & Halbert, S. E. Asian citrus psyllid, *Diaphorina citri*, vector of citrus huanglongbing disease. *Entomol. Exp. Appl.* **146**, 207–223 (2013).
79. Huang, F. *et al.* *Diaporthe* species occurring on citrus in China. *Fungal Divers.* **61**, 237–250 (2013).
80. Moghbeli Gharaei, A., Ziaaddini, M., Jalali, M. A. & Michaud, J. P. Sex-specific responses of Asian citrus psyllid to volatiles of conspecific and host-plant origin. *J. Appl. Entomol.* **138**, 500–509 (2014).
81. Althoff, D. M., Segreaves, K. A. & Johnson, M. T. Testing for coevolutionary diversification: linking pattern with process. *Trends. Ecol. Evol.* **29**, 82–89 (2014).
82. Servedio, M. R. Geography, assortative mating, and the effects of sexual selection on speciation with gene flow. *Evol. Appl.* **9**, 91–102 (2016).
83. Ferrari, J., Darby, A. C., Daniell, T. J., Godfray, H. C. J. & Douglas, A. E. Linking the bacterial community in pea aphids with host-plant use and natural enemy resistance. *Ecol. Entomol.* **29**, 60–65 (2004).
84. Yang, X. *et al.* Genetic diversity and phylogenetic relationships of citron (*Citrus medica* L.) and its relatives in southwest China. *Tree. Genet. Genomes.* **11**, 129 (2015).
85. Beck, S. D. & Reese, J. C. Insect-plant interactions: nutrition and metabolism. *Recent Advances in Phytochemistry* **10**, 41–92 (1976).
86. Russell, D. N., Qureshi, J. A., Halbert, S. E. & Stansly, P. A. Host suitability of citrus and *Zanthoxylum* spp. for *Leuronota fagarae* Burckhardt and *Diaphorina citri* Kuwayama (Hemiptera: Psylloidea). *Fla. Entomol.* **97**, 1481–1492 (2014).
87. Liu, Q. *et al.* Temperature, precipitation, and insolation effects on autumn vegetation phenology in temperate China. *Global Change. Biol.* **22**, 644–655 (2016).
88. Sunnucks, P. & Hales, D. F. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Mol. Biol. Evol.* **13**, 510–524 (1996).
89. Li, Q. & Wan, J. M. SSR hunter: development of local searching software for SSR sites. *Hereditas* **27**, 808–810 (2005).
90. Rozen, S. & Skaletsky, H. Primer 3 on the WWW for general users and for biologist programmers. *Methods. Mol. Biol.* **6**, 365–386 (1999).
91. de Arruda, M. P., Gonçalves, E. C., Schneider, M. P. C., da Silva, A. L. D. C. & Morielle-Versute, E. An alternative genotyping method using dye-labeled universal primer to reduce unspecific amplifications. *Mol. Biol. Rep.* **37**, 2031–2036 (2010).
92. Peakall, R. O. D. & Smouse, P. E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes.* **6**, 288–295 (2006).
93. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes.* **4**, 535–538 (2004).
94. Raymond, M. & Rousset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Heredity* **86**, 248–249 (1995).
95. Goudet, J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* **86**, 485–486 (1995).
96. Cavalli-Sforza, L. L. & Edwards, A. W. Phylogenetic analysis. Models and estimation procedures. *Am. J. Hum. Genet.* **19**, 233 (1967).
97. Felsenstein, J. PHYLIP: phylogenetic inference package, version 3.5 c. (1993).
98. Page, R. D. Treeview: an application to display phylogenetic trees on personal computers. *Bioinformatics* **12**, 357 (1996).
99. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491 (1992).
100. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
101. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids. Res.* **25**, 4876–4888 (1997).
102. Hall, T. A. BioEdit: a user-friendly biological sequence alignment program for Windows 95/98/NT. *Nucleic. Acids. Symposium.* **41**, 95–98 (1999).
103. Nei, M. & Jin, L. Variances of the average numbers of nucleotide substitutions within and between populations. *Mol. Biol. Evol.* **6**, 290–300 (1989).
104. Librado, P. & Rozas, J. Dnasvp5: a software for comprehensive analysis of dna polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
105. Hasegawa, M., Kishino, H. & Yano, T. A. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).
106. Slatkin, M. Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* **47**, 264–279 (1993).
107. Mantel, N. Ranking procedures for arbitrarily restricted observation. *Biometrics* **23**, 65–78 (1967).
108. 114.Swofford, D. L. PAUP*: Phylogenetic Analysis using Parsimony (*and OtherMethods): version 4.0b10. Sinauer, Sunderland. (2002).
109. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
110. Leigh, J. W. & Bryant, D. popart: full-feature software for haplotype network construction. *Methods. Ecol. Evol.* **6**, 1110–1116 (2015).
111. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).

112. Fu, Y. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).
113. Ramos-Onsins, S. E. & Rozas, J. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* **19**, 2092–2100 (2002).
114. Schuelke, M. An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* **18**, 233–234 (2000).

Acknowledgements

This work was supported by National Key R & D Program of China (No. 2017YFD0202000), the earmarked fund for the China Agricultural Research System (No.CARS-26) to HYZ. WHW is funded by MBIE via New Zealand – China Water Research Centre. We thank all citrus experimental stations of the China Agricultural Research System (No. CARS-26) for collecting the Asian citrus psyllid.

Author Contributions

L.X.M. and H.Y.Z. designed the work, conducted the data analysis, and wrote the paper. L.X.M. carried out fieldwork and performed laboratory and data analyses. Y.M.W. and W.H.W. read, edited and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-19533-5>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018