Research

Identification of the Population Structure of *Myzus persicae* (Hemiptera: Aphididae) on Peach Trees in China Using Microsatellites

Jie Li[1], Jinjun Cao[1], Jianqun Niu, Xiaoxia Liu, and Qingwen Zhang[2]

Department of Entomology, College of Agriculture and Biotechnology, China Agricultural University, Beijing, China
[1]Jie Li and Jinjun Cao are co-first authors; they contributed equally to the work.
[2]Corresponding author, e-mail: zhangqingwen@263.net

Subject Editor: Sara Goodacre

J. Insect Sci. *(2015)* 15(1): 73; DOI: 10.1093/jisesa/iev026

ABSTRACT. In this study, we characterized the genetic structure of *Myzus persicae* (Sulzer) (Hemiptera: Aphididae) populations in China using microsatellites. We expected that these data will reveal the genetic relationships among various populations of *M. persicae* and will be of value in the development of better methods for pest control. Four hundred sixty individuals from 23 areas over 13 provinces were collected in the early spring of 2010, all from their primary host, *Prunus persicae*. The markers analyzed were highly polymorphic, as demonstrated by the expected heterozygosity value ($H_e = 0.861$) and the Polymorphism Information Content (PIC = 0.847), which indicated that *M. persicae* maintains a high level of genetic diversity. Analysis of molecular variance revealed an intermediate level of population differentiation among *M. persicae* populations ($F_{ST} = 0.1215$). Geographic isolation existed among these populations, and, consequently, the genetic structure of the populations was split into a southern group and a northern group divided by the Yangtse River.

Key Words: *Myzus persicae*, microsatellite, genetic diversity, genetic structure

Myzus persicae (Sulzer) (Hemiptera: Aphididae) is presumed to be of Asian origin, but it is globally distributed today and represents an economically important pest (Blackman and Eastop 2000). *M. persicae* is a polyphagous aphid species that feeds on over 400 plant species in 40 different families and is the most important vector of plant viruses, such as Potato leafroll virus (Van Emden et al. 1969, Blackman and Eastop 1984, Van den Heuvel et al. 1994).

Like many aphid species, *M. persicae* has a complex and variable life cycle. Zhang and Zhong (1983) reported that in China, *M. persicae* normally reproduces by cyclical parthenogenesis on its primary host, peach, *Prunus persicae* (L.) Batsch. The organism undergoes several generations of apomictic parthenogenesis followed by a single sexual generation. Mating takes place on the primary host, where the eggs are laid and undergo diapause over the winter. Parthenogenetic females hatch in the spring, and their descendants disperse to secondary host plants, where they undergo many parthenogenetic generations (Blackman and Eastop 2000).

Microsatellites are a class of codominant and hypervariable genetic markers (Jarne and Lagoda 1996) and have considerable potential for analyzing the population structures of aphid species (Llewellyn et al. 2003, Papura et al. 2003, Simon et al. 2003, Vorwerk and Forneck 2006, Cao et al. 2012). Fourteen microsatellite loci in *M. persicae* were separately described by Sloane et al. (2001). These loci have been used to study population structures by screening for genetic diversity (Wilson et al. 2002, Fenton et al. 1998, 2003, Guillemaud et al. 2003, Vorburger et al. 2003, Fuentes-Contreras et al. 2004, Vorburger 2006, John et al. 2009). Therefore, we expected that these loci could be used to analyze population structures in Chinese *M. persicae*.

Although *M. persicae* have long been recognized as one of the most important agricultural pests in China, little has been reported about its genetic diversity on peaches. Previous studies have focused on a few areas, different methods, and genetic variation in aphid populations on different host plants (Yang and Zhang 1999, Han et al. 2009, Liu et al. 2010). Furthermore, documentation of genetic structure and its distribution across the landscape is important to understanding the role of genetic variation in the success or failure of an invasion (Novak and Mack 2005, Lavergne and Molofsky 2007). Therefore, to obtain a thorough understanding of the *M. persicae* population structure on peaches and to better understand population phylogenetic relationship for forecasting and improving pest control, we analyzed the genetic diversity of various geographical populations of *M. persicae* on peaches using microsatellite markers.

Materials and Methods

Sampling Strategy. A total of 460 parthenogenetic aphids were collected from 23 areas distributed over 13 provinces in China mainly in early spring (late April to July) of 2010 (Fig. 1; Table 1). Because of the cyclical parthenogenetic life cycle of the aphid, the spring population has a large amount of variation that may decrease in later seasons due to natural selection, artificial selection, or drift (Guillemaud et al. 2003). Therefore, to fully reflect the genetic diversity among populations, sampling in early spring is necessary. Each aphid was sampled from different plants of peach, which were separated from each other by more than 50 m to minimize the risk of collecting organisms of the same clone. In most cases, the distance between any two populations was greater than 50 km. Samples from natural populations were preserved in 95% ethanol and stored at −20°C. Within each population, 20 individuals were sampled for analysis (Table 1).

DNA Isolation and PCR Assay. Total genomic DNA was extracted from a single aphid using the cetyltrimethylammonium bromide method (Doyle and Doyle 1987). An UV-visible spectrophotometer (Thermo Scientific NanoDrop 2000c) was used to detect the content and purity of the extracted DNA. All DNA samples were diluted to 40 ng/μl and stored at −20°C for future use.

Seven polymorphic microsatellite loci were used in this study (Table 2). Five of these, M35, M40, M49, M63, and M86, were previously identified in an Australian clonal lineage of *M. persicae* (Sloane et al. 2001). The additional two loci, myz2 and myz25, were identified in a British clone of *M. persicae* (G. Malarky, unpublished data). The full details of microsatellite testing and amplification were published in the report by Sloane et al. (2001).

Data Analysis. Our results were interpreted using GeneMapper software (version 4.0) (Applied Biosystems, Foster City, CA), which

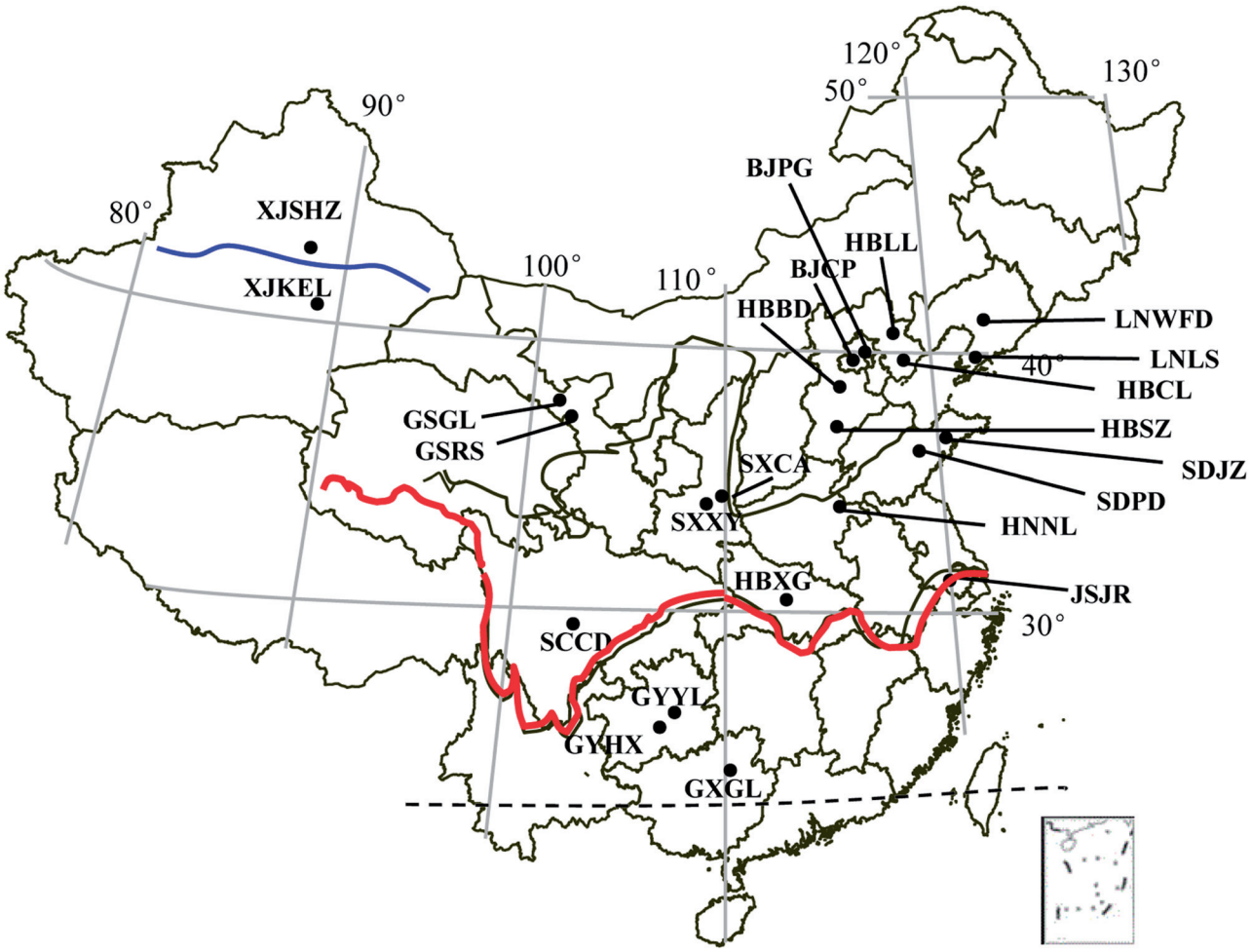

Fig. 1. Locations of the 23 populations sampled within China. Population codes are the same as in Table 1. Red line represents the Yagtse River. Blue line represents the Tian Shan Mountain.

calculates allele sizes at each microsatellite locus compared to the standard size. For each population and loci, the number of alleles (N_a), the effective number of alleles (N_e), the observed heterozygosity (H_o), and the expected heterozygosity (H_e) were calculated by PopGene32 (version 1.31) (University of Alberta, Calgary, Canada) (Yeh et al. 1999). The Polymorphism Information Content (PIC) for each microsatellite locus was calculated according to Bostein et al. (1980). Arlequin version 3.11 (University of Bern, Bern, Switzerland) (Excoffier et al. 2005) was used to analyze Hardy–Weinberg equilibrium (HWE), linkage disequilibrium, and to calculate F_{IS} values (Weir and Cockerham 1984).

Pairwise genetic distances and F_{ST} values were calculated in Arlequin version 3.11. We also used this software for analysis of molecular variance (AMOVA) (Ordonez and Kolmer 2007).

The principal coordinate analysis (PCA) on a genetic distance matrix (Nei's unbiased genetic distance; Nei 1972, 1978) was analyzed in GENEAlEX V6.1 (Australian National University, Canberra, Australia) (Peakall and Smouse 2006) for all populations. The genetic structure and an estimate of the most likely number of clusters (gene pools) were inferred by the Bayesian clustering method using STRUCTURE 2.0 (Oxford University, Oxford, UK) (Pritchard et al. 2000). The data set was analyzed using the admixture and uncorrelated allele frequencies models and K values 1–14 without incorporating population information. Four independent runs for each K were conducted with 100,000 iterations after a burn-in period of 10,000 iterations in each run.

A phylogenetic tree, based on Nei's unbiased genetic distance using the UPGMA clustering method (an unweighted pair-group method using arithmetical averages), was constructed using the PHYLIP 3.66 software (University of Washington, Seattle, WA) (Felsenstein 2004).

For detecting an association between genetic and geographic distances, Mantel's test was implemented with 1,000 permutations using the program IBD version 1.5.2 (San Diego State University, San Diego, CA) (Bohonak 2002). The value of $F_{ST}/(1 - F_{ST})$ was calculated using Arlequin version 3.11, geographical distance between populations was calculated according to the latitude and longitude of the location of each population, and the natural logarithm of distance was used as the measure of geographical distance to reduce error (Sokal and Rohlf 1995). Alleles in Space (http://www.marksgeneticsoftware.net/) (Miller 2005) package including the Allelic Aggregation Index Analysis (AAIA) tested for nonrandom patterns of allele phenotype diversity across the landscape.

Results

Microsatellite Diversity. The seven microsatellite loci were highly polymorphic, with between 22 and 45 alleles per locus (Table 2) and an average of 29. The observed heterozygosity (H_o) was an average of 0.688 (range: 0.513–0.815). The expected heterozygosity (H_e) was an average of 0.861 (range: 0.767–0.939). In tests of HWE, we found that with the exception of M40 ($F_{IS} = -0.088$) and M63 ($F_{IS} = -0.002$), the loci exhibited a deficiency of heterozygosity ($F_{IS} > 0$). PIC for each

Table 1. Population information and genetic variability estimates based on data from 7 microsatellite loci in 23 populations of *M. persicae*

Code	Collection site	Latitude (N)/longitude (E)	*n*	N_a	N_e	H_o	H_e	F_{IS}
GYHX	Huaxi, Guiyang	26° 25′ 29.20"/106° 40′ 19.75"	20	9.571	4.176	0.493	0.690	0.291*
GYYL	Yongle, Guiyang	26° 35′ 52.76"/106° 52′ 27.22"	20	7.857	4.443	0.400	0.586	0.322*
GXGL	Guilin, Guangxi	24° 59′ 21.65"/110° 51′ 0.40"	20	10.571	4.519	0.607	0.774	0.220*
SCCD	Chengdu, Sichuan	30° 32′ 29.67"/104° 18′ 45.67"	20	7.571	4.031	0.729	0.733	0.006*
HBXG	Xiaogan, Hubei	31° 1′ 35.77"/114° 5′ 43.01"	20	9.571	5.616	0.664	0.827	0.201*
SXXY	Xianyang, Shanxi	34° 19′ 44.27"/108° 44′ 32.86"	20	8.286	4.720	0.557	0.774	0.285*
SXCA	Changan, Shanxi	34° 3′ 22.59"/109° 3′ 55.70"	20	8.571	5.035	0.671	0.776	0.137*
JSJR	Jurong, Jiangsu	31° 46′ 9.37"/119° 11′ 7.04"	20	8.143	4.404	0.514	0.571	0.102*
SDJZ	Jiaozhou, Shandong	36° 22′ 3.69"/119° 57′ 27.18"	20	8.286	4.555	0.593	0.755	0.219
SDPD	Pingdu, Shandong	36° 48′ 21.65"/119° 36′ 56.19"	20	8.000	5.539	0.757	0.812	0.070
HNNL	Ningling, Henan	34° 29′ 54.75"/115° 18′ 24.84"	20	8.143	4.052	0.664	0.728	0.089
HBSZ	Shenzhou, Hebei	37° 59′ 34.74"/115° 31′ 54.01"	20	8.143	5.395	0.771	0.801	0.037*
HBBD	Baoding, Hebei	38° 50′ 31.03"/115° 7′ 59.15"	20	9.857	5.926	0.736	0.801	0.083
HBCL	Changli, Hebei	39° 42′ 46.14"/119° 9′ 45.73"	20	8.714	5.255	0.771	0.793	0.028
HBLL	Lulong, Hebei	39° 53′ 22.34"/118° 54′ 58.73"	20	9.286	5.023	0.757	0.786	0.038*
BJCP	Changping, Beijing	40° 17′ 49.50"/116° 13′ 32.70"	20	9.857	5.974	0.700	0.819	0.149
BJPG	Pinggu, Beijing	40° 8′ 3.43"/117° 1′ 19.28"	20	9.571	5.879	0.736	0.813	0.097*
GSGL	Gaolan, Gansu	36° 20′ 0.83"/103° 56′ 49.50"	20	10.857	6.587	0.786	0.841	0.068
GSRS	Renshou, Gansu	36° 4′ 34.55"/103° 45′ 29.90"	20	10.857	6.866	0.800	0.848	0.058
LNLS	Lvshun, Liaoning	38° 48′ 53.90"/121° 13′ 2.67"	20	4.429	2.391	0.843	0.579	−0.473
LNWFD	Wafangdian, Liaoning	39° 37′ 37.61"/121° 58′ 46.57"	20	5.286	3.434	0.836	0.719	−0.168
XJSHZ	Shihezi, Xinjiang	44° 17′ 39.43"/85° 51′ 14.14"	20	9.429	5.244	0.800	0.795	−0.007
XJKEL	Kuerle, Xinjiang	41° 45′ 39.51"/86° 9′ 49.90"	20	10.857	6.074	0.629	0.821	0.239*
Means			20	8.770	5.006	0.688	0.758	0.011

n, number of individuals per population; N_a, observed number of alleles per locus; N_e, effective number of alleles per locus; observed (H_o) and expected heterozygosity (H_e). F_{IS}, inbreeding coefficient.

*Significance at the 5% nominal level.

Table 2. Microsatellite primers and genetic variation among seven microsatellite loci of *M. persicae* in China

Locus	Repeat motif	Primer sequence (5′–3′)	T_a (°C)	N_a	N_e	H_o	H_e	F_{IS}	PIC
M2	(GA)30	H-TGGCGAGAGAGAAGACCTGC TCGGAAGACAGAGACATCGAGA	58	23	6.468	0.694	0.846	0.076	0.831
M25	(AG)24	F-GAATCTGGAGAGCGGTTAATGC AACCCATCTCACTCGTCAGCC	55	23	4.278	0.657	0.767	0.010	0.742
M35	(AT)9-(AC)22	T-GGCAATAAAGATTAGCGATG TGTGTGTATAGATAGGATTTGTG	55	22	9.285	0.513	0.893	0.286	0.883
M40	(AC)17	H-ACACGCATACAAGAATAGGG AGAGGAGGCAGAGGTCAAAC	55	23	4.967	0.741	0.800	−0.088	0.779
M49	(AC)31	F-CCCATACATACCTCCAAGAC AGAGAGAAAATAGGTTCGTG	49	45	16.155	0.772	0.939	0.033	0.935
M63	(AC)29	T-GATTATGGTGCTCGGTGG GCGGTTTTCTTTGTATTTTCG	49	40	9.528	0.815	0.896	−0.002	0.887
M86X	(CA)23	H-TCCACTAAGACCTCAAACAC ATTTATTATGTCGTTCCGCC	55	27	8.567	0.622	0.884	0.173	0.872
Mean				29	8.464	0.688	0.861	0.070	0.847

The fluorophore is attached to the 5′-end of the forward primer. F, H, and T represent three different fluorophores. F (FAM) is blue fluorescence, H (HEX) is green fluorescence, and T (TAMRA) is yellow fluorescence. T_a, annealing temperature; N_a, number of alleles; N_e, Effective number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; F_{IS}, inbreeding coefficient; PIC, polymorphic information content.

locus had a minimum value of 0.742 for the locus M25 and a maximum value of 0.935 for the locus M49 (Table 2).

Within Population Genetic Diversity. The observed number of alleles per locus for each population ranged from 4.429 (LNLS) to 10.857 (GSGL, GSRS, XJKEL), with an average of 8.770 (Table 1). The average effective number of alleles per locus was 5.006, with a maximum of 6.866 (GSRS) and a minimum of 2.391 (LNLS) (Table 1). The observed (H_o) and expected heterozygosity (H_e) values ranged from 0.400 (GYYL) to 0.843 (LNLS) (average = 0.688) and from 0.586 (GYYL) to 0.848 (GSRS) (average = 0.758), respectively (Table 1). F_{IS} values ranged from −0.473 (LNLS) to 0.323 (GYYL) (average = 0.011), and only three populations had negative F_{IS} values (LNLS, LNWFD, XJSHZ) (Table 1).

Genetic Differentiation Among Populations. The PCA was able to validate the differentiation that was apparent in the pairwise F_{ST} values (Fig. 2). The first two PC axes cumulatively accounted for 63.72% of the total variation. PC axis 1 accounted for 49.00% of the variation; along the first axis, populations from southern China (SCCD, GXGL, GYYL, GYHX and JSJR) were distinct from the majority of the collected samples. Populations were less divergent along PC axis 2, which accounted for 14.72% of the total variation. Along the second axis, a slight differentiation of LNLS, JSJR and XJKEL from other populations was observed.

With the exception of nine pairwise comparisons, all the tests for pairwise genetic differentiation among populations were significant (Table 3). Pairwise F_{ST} values ranged from 0.002 (between the GYHX and GYYL populations) to 0.410 (between the LNLS and JSJR populations), suggesting a low genetic differentiation among GYHX and GYYL and a high genetic differentiation among LNLS and JSJR. The genetic distance ranged from 0.061 (between GYHX and GYYL) to 2.873 (between LNLS and JSJR) reinforced the result above.

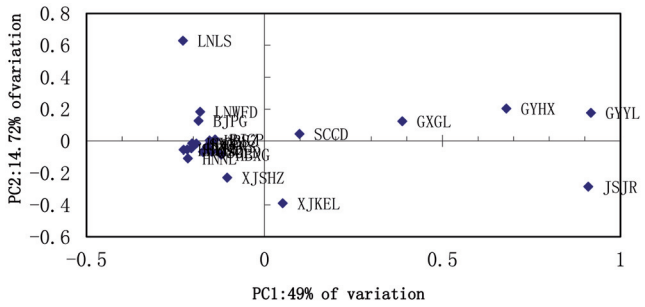

Fig. 2. PCA among 23 *M. persicae* populations.

From the AMOVA, we found that 12.15% ($P < 0.05$) of the genetic variation was between populations, and the remaining 79.46% ($P < 0.05$) of the genetic variation was between individuals, representing the main source of genetic variation. Genetic differentiation estimates reached an $F_{ST} = 0.1215$, suggesting moderate genetic differentiation (Table 4).

STRUCTURE analysis estimated the log likelihood and K values (Supp. Fig. 1 [online only]) (Evanno et al. 2005). The most probable division is $K = 3$, identifying three identifiable genetic clusters among the 23 populations (Fig. 3b). Each regional subspecies formed a unique set of clusters.

Cluster 1 (M1) contained almost all the lineages from GYHX, GYYL, and JSJR. Cluster 2 (M2) is characterized by the lineages collected from SCCD, XJSHZ, XJKEL, and a large portion of GXGL, HNNL, GSGL, and GSRS were assigned to Cluster 2. All the other populations were assigned to Cluster 3 (M3) (Fig. 3b).

A similar pattern of differentiation among populations was constructed with Nei's unbiased genetic distance using the UPGMA clustering method (Fig. 3a). The phylogenetic tree showed that four southern Chinese populations (GYYL, GYHX, GXGL, and JSJR) clustered together and then clustered with other populations from north and northwest China.

Isolation by Distance Among Populations. The geographical distance between any two populations ranged from 27.55 to 3,209.23 km. Most of the distances between two populations were greater than 50 km except the distance between GYHX and GYYL, SXXY and SXCA, HBCL and HBLL, GSGL and GSRS. For the Mantel test, the value of $F_{ST}/(1 - F_{ST})$ was calculated using Arlequin 3.11; the natural logarithm of distance was used as the measure of geographical distance to reduce error. A pattern of isolation by distance (IBD) was evident by the positive regression between linearized F_{ST} and geographic distance ($r = 0.2864$, $P = 0.0110$) (Fig. 4). In testing the microsatellite allelic data for spatial patterns due to gene flow, the AAIA test for nonrandom genotypic patterns among populations displayed non-significant evidence of a spatial pattern of allele distribution (spatial aggregation of sampled points: $r = 0.7010$, $P = 0.2210$). For the genetic landscape interpolation, the genetic distance between populations from the northwest to the southern existed big variations, whereas from northeast to southeast, the genetic distance tends to be stable although with small fluctuations (Fig. 5).

Discussion

All seven microsatellite loci showed a high polymorphism: N_a ranged from 22 to 45, average PIC was 0.847(0.5), and locus M49 showed the greatest variability. This phenomenon was similar to many previous studies (Wilson et al. 2002, Fuentes-Contreras et al. 2004, John et al. 2009) and illustrated that sample collection (Maudet et al. 2002) and loci selection were at a sensible level.

In our study, high levels of polymorphism were identified in all loci within each population. According to Zhang and Zhong (1983), *M. persicae* in China is cyclically parthenogenetic (holocyclic); therefore, a high genetic diversity, a heterozygote deficit, and a positive F_{IS} value were expected in its primary host the peach. As Wilson et al. (2003) proved, a high genetic diversity may be the result of genetic recombination which has taken place in mating once a year. Delmotte et al. (2002) examined how reproductive mode shapes genetic structure of sexual (cyclically parthenogenetic) and asexual (obligately parthenogenetic) populations of the aphid *Rhopalosiphum padi* by comparing microsatellite and allozyme data sets. Microsatellites indicated that sexual populations have high allelic polymorphism and heterozygote deficits (possibly because of population subdivision, inbreeding or selection). Wilson et al. (2002) considered that heterozygote excess has been found in asexual lineages is attributed either to ancient loss of sexuality and the consequence of accumulated mutations or to a hybrid origin. We think heterozygosity in aphid may be related to reproduction mode and molecular inheritance marker method. In our study, the H_o values in most populations were smaller than the H_e values, with the exception of the LNLS, LNWFD, and XJSHZ populations; and the FIS values of these three populations were negative, especially of LNLS, which was strongly negative (−0.473). Many explanations for this phenomenon have been discussed in previous studies, including selection, clonal expansion, the Wahlund effect, inbreeding, and other population effects (Fenton et al. 2003, John et al. 2009). Within the category of selection, natural selection (e.g., selective pressures from temperature, rainfall, and natural predators) and human selection (e.g., insecticides, plant trade, and transport) have been commonly used to explain our observed phenomenon in different aphid species (Delmotte et al. 2002, Fenton et al. 2003, Vorburger 2006). Wahlund effect of sampling from distinct gene pools in the same population may also contribute to the homozygote excess (Fenton et al. 2003, John et al. 2009).

The AMOVA revealed an intermediate level ($0.05 < F_{ST} \leq 0.15$; Wright 1978) of population differentiation ($F_{ST} = 0.1215$) among populations of *M. persicae* in China. The results of the STRUCTURE analysis also revealed a moderate level of differentiation among the populations of *M. persicae*. Similarly, moderate population genetic differentiation was suggested by the UPMGA tree and the PCA test too. Meanwhile, the high within individual and intermediate inter-population genetic diversity of *M. persicae* reflect some gene flow among populations. The inferred gene flow among some populations based on the F_{ST} values revealed were high (N_m values ranging from 0.360 to 124.75) (data not shown).

The pairwise F_{ST} was not significant over a small geographic distance (<50 km) and was similar to the results observed by Guillemaud et al. (2003) (<60 km). As an explanation for this phenomenon, Guillemaud et al. discussed previous findings that interpopulation differentiation and aphid species had a clear relationship and that large differences in migration capacities existed between aphid species. In addition, the effects of natural forces on migration should be taken into account. Our primary research demonstrated that southeast monsoons originating from the Pacific Ocean in May strongly influence the migration of the *Spiraea* aphid (Cao et al. 2012). Similarly, the influence of topographic factors (e.g., mountains and rivers) of population genetic differentiation should be considered.

The Yangtse River divides China's geography and distinct climates into northern and southern regions. As we know, geographic isolation, such as mountains and rivers, is a major factor contributing to genetic differentiation. They can lead to different geographic populations cannot freely exchanged. The UPMGA phylogenetic tree and the PCA demonstrated a clear divide between a southern and northern group separated by the Yangtse River. In the northern group, the LNLS population, which is seaside, was widely separated from other populations. Compared with XJSHZ population, XJKEL population is further from other northern population. Because, Tian Shan Mountain is the boundary of northern and southern Xinjiang, XJSHZ population in northern Xinjiang is convenient to exchange to other populations. This finding illustrated that the geographic isolation between northern and southern populations as well as differences in the climate between the north and

Table 3. F_{ST} (above diagonal) and Nei's unbiased genetic distance (below diagonal) among 23 populations of *M. persicae* on peaches in China

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1		0.002*	0.044	0.125	0.169	0.171	0.186	0.193	0.194	0.169	0.209	0.180	0.174	0.177	0.174	0.153	0.166	0.163	0.158	0.280	0.210	0.203	0.200
2	0.061		0.100	0.204	0.249	0.254	0.264	0.202	0.270	0.245	0.295	0.258	0.260	0.256	0.261	0.236	0.252	0.242	0.238	0.368	0.297	0.275	0.261
3	0.189	0.283		0.058	0.096	0.104	0.123	0.213	0.127	0.108	0.142	0.100	0.104	0.110	0.103	0.087	0.093	0.096	0.095	0.228	0.133	0.138	0.144
4	0.490	0.714	0.281		0.106	0.085	0.105	0.293	0.117	0.082	0.125	0.070	0.070	0.083	0.067	0.064	0.092	0.076	0.077	0.232	0.153	0.118	0.155
5	1.058	1.591	0.647	0.611		0.035	0.055	0.254	0.060	0.050	0.069	0.061	0.034	0.054	0.037	0.043	0.034	0.033	0.028	0.168	0.077	0.072	0.108
6	0.887	1.296	0.591	0.407	0.245		0.017*	0.295	0.027*	0.040	0.061	0.052	0.014*	0.029	0.019	0.021	0.033	0.029	0.025	0.181	0.073	0.099	0.134
7	1.043	1.455	0.743	0.517	0.356	0.143		0.308	0.029	0.036	0.113	0.057	0.049	0.037	0.049	0.033	0.052	0.045	0.045	0.200	0.089	0.116	0.143
8	0.555	0.462	0.819	1.514	1.547	1.968	2.484		0.317	0.283	0.329	0.293	0.297	0.301	0.296	0.274	0.288	0.276	0.270	0.410	0.327	0.293	0.267
9	1.037	1.418	0.717	0.555	0.352	0.170	0.179	2.449		0.033	0.083	0.073	0.039	0.036	0.044	0.042	0.049	0.055	0.052	0.207	0.098	0.103	0.149
10	1.000	1.396	0.713	0.431	0.381	0.261	0.243	2.160	0.210		0.074	0.058	0.034	0.036	0.039	0.028	0.054	0.042	0.041	0.179	0.102	0.094	0.114
11	1.093	1.646	0.757	0.553	0.363	0.288	0.556	2.435	0.368	0.375		0.109	0.022	0.084	0.039	0.074	0.073	0.062	0.055	0.211	0.138	0.124	0.137
12	1.074	1.534	0.618	0.356	0.439	0.316	0.340	2.359	0.401	0.392	0.572		0.052	0.040	0.049	0.040	0.036	0.034	0.043	0.186	0.070	0.092	0.128
13	1.012	1.587	0.649	0.354	0.270	0.142	0.301	2.548	0.234	0.258	0.144	0.345		0.024	0.003*	0.020	0.030	0.007*	0.005*	0.158	0.093	0.077	0.108
14	1.016	1.455	0.680	0.416	0.374	0.199	0.235	2.584	0.217	0.262	0.413	0.270	0.193		0.020	0.012*	0.041	0.028	0.039	0.172	0.093	0.097	0.121
15	0.959	1.481	0.612	0.329	0.269	0.156	0.289	2.180	0.247	0.271	0.206	0.311	0.100	0.171		0.019	0.033	0.016	0.022	0.173	0.104	0.096	0.124
16	0.853	1.301	0.555	0.337	0.344	0.176	0.233	1.962	0.257	0.239	0.385	0.295	0.189	0.151	0.178		0.029	0.031	0.028	0.153	0.087	0.084	0.111
17	0.965	1.530	0.590	0.489	0.282	0.227	0.324	2.388	0.288	0.388	0.372	0.266	0.240	0.285	0.238	0.248		0.030	0.024	0.122	0.056	0.085	0.129
18	1.043	1.565	0.683	0.427	0.308	0.220	0.311	2.420	0.338	0.347	0.334	0.280	0.137	0.236	0.170	0.283	0.271		0.003*	0.159	0.076	0.058	0.096
19	1.009	1.538	0.691	0.436	0.283	0.204	0.314	2.284	0.326	0.348	0.300	0.342	0.129	0.302	0.200	0.272	0.242	0.109		0.141	0.088	0.064	0.097
20	1.162	1.712	0.970	0.886	0.626	0.631	0.757	2.873	0.756	0.680	0.725	0.708	0.540	0.605	0.602	0.528	0.371	0.587	0.483		0.184	0.213	0.248
21	1.060	1.606	0.670	0.715	0.391	0.328	0.405	2.178	0.430	0.524	0.605	0.335	0.455	0.449	0.499	0.443	0.278	0.397	0.477	0.568		0.142	0.179
22	1.383	1.885	0.957	0.636	0.505	0.594	0.742	2.247	0.579	0.645	0.671	0.607	0.492	0.622	0.605	0.580	0.579	0.423	0.477	0.901	0.784		0.102
23	1.524	1.804	1.175	1.090	0.953	1.029	1.167	1.770	1.136	0.959	0.851	1.089	0.835	0.951	0.958	0.949	1.181	0.853	0.895	1.463	1.370	0.752	

Nei's genetic distance (below diagonal) and F_{ST} (above diagonal); 1: GYHX; 2: GYYL; 3: GXGL; 4: SCCD; 5: HBXG; 6: SXXY; 7: SXCA; 8: JSJR; 9:SDJZ; 10: SDPD; 11: HNNL; 12:HBSZ; 13: HBBD; 14: HBCL; 15: HBLL; 16: BJCP; 17:BJPG; 18: GSGL; 19:GSRS; 20: LNLS; 21: LNWFD; 22: XJSHZ; 23: XJKEL.

*Nonsignificant values ($P > 0.05$).

Table 4. AMOVA of microsatellites in 23 *M. persicae* populations

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Fixation indices	*P*
Among populations	22	387.972	0.3680	12.15	$F_{ST} = 0.1215$	0.0000
Among individuals within populations	437	1273.775	0.2542	8.39	$F_{IS} = 0.0955$	0.0000
Within individuals	460	1107	2.4065	79.46	$F_{IT} = 0.2054$	0.0000
Total	919	2768.747	3.02868			

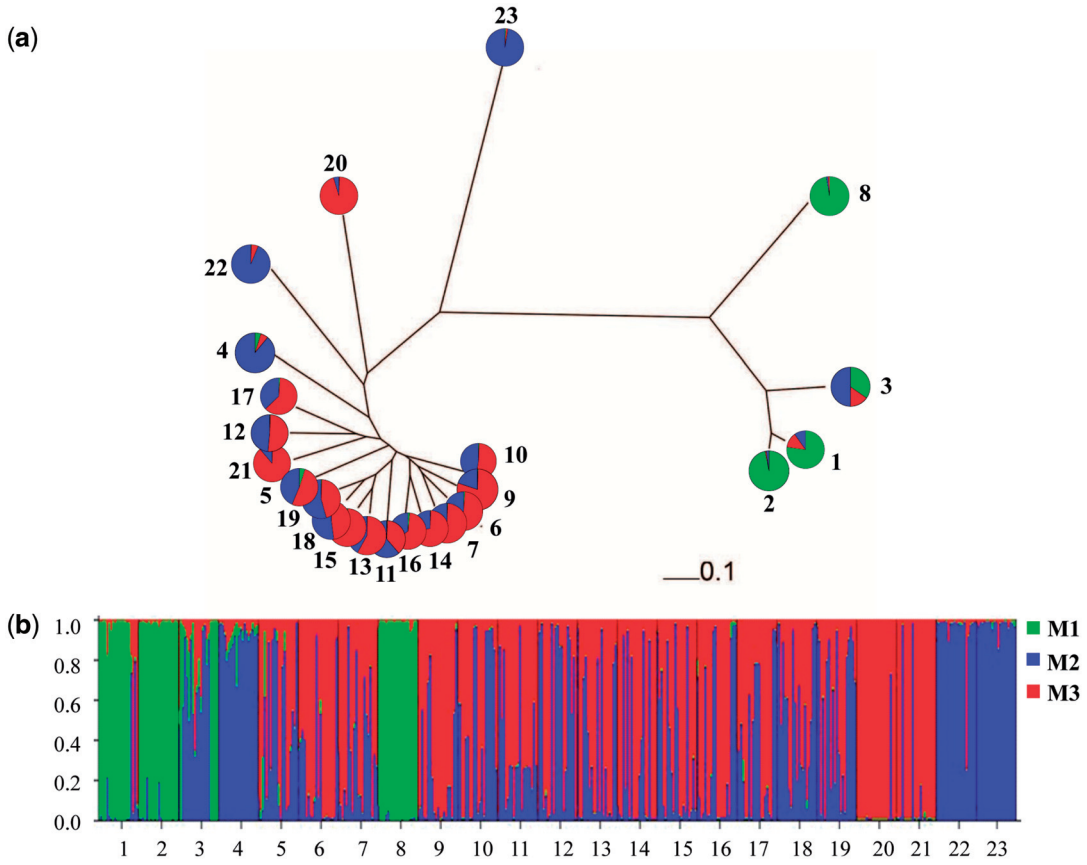

Fig. 3. (a) Unrooted UPGMA consensus tree constructed from Nei's unbiased genetic distance depicting the relationships of 23 populations of *M. persicae*. (b) A population assignment test using STRUCTURE (version 2.3.2) software based on eight microsatellite loci produced from different groups ($K = 3$). The vertical lines are broken into colored segments showing the proportion of each individual assigned to each of the inferred K. Geographic regions from which the populations belong appear along the x-axis. 1: GYHX; 2: GYYL; 3: GXGL; 4: SCCD; 5: HBXG; 6: SXXY; 7: SXCA; 8: JSJR; 9: SDJZ; 10: SDPD; 11: HNNL; 12: HBSZ; 13: HBBD; 14: HBCL; 15: HBLL; 16: BJCP; 17: BJPG; 18: GSGL; 19: GSRS; 20: LNLS; 21: LNWFD; 22: XJSHZ; 23: XJKEL.

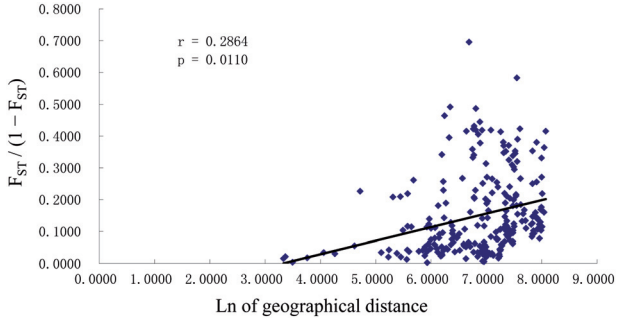

Fig. 4. Isolation-by-distance plot of $F_{ST}/(1 - F_{ST})$ versus the natural log of geographical distance (km). The solid line represents the best-fit linear regression based on all points.

south accelerated genetic differentiation among populations of *M. persicae* to some extent.

The Bayesian clustering and admixture analysis indicated that each regional subspecies formed a unique set of clusters (or gene pools). The genetic variation partitioned into three clusters. Those indicated that *M. persicae* populations were moving to three different evolution directions. For those phenomena, migration of aphids should firstly be considered. In addition, to some extent, this effect can be caused by anthropogenic activity (John et al. 2009), such as the trading of plants. *M. persicae* is ideally suit for this, as its primary host, the peach tree, has been spread throughout the nation and has had an immediate impact on the evolution of pest populations.

In this study, there was an obscure relationship between genetic and geographic distance ($r = 0.2864$, $P = 0.0110$). However, our previous research identified a significant correlation between genetic and

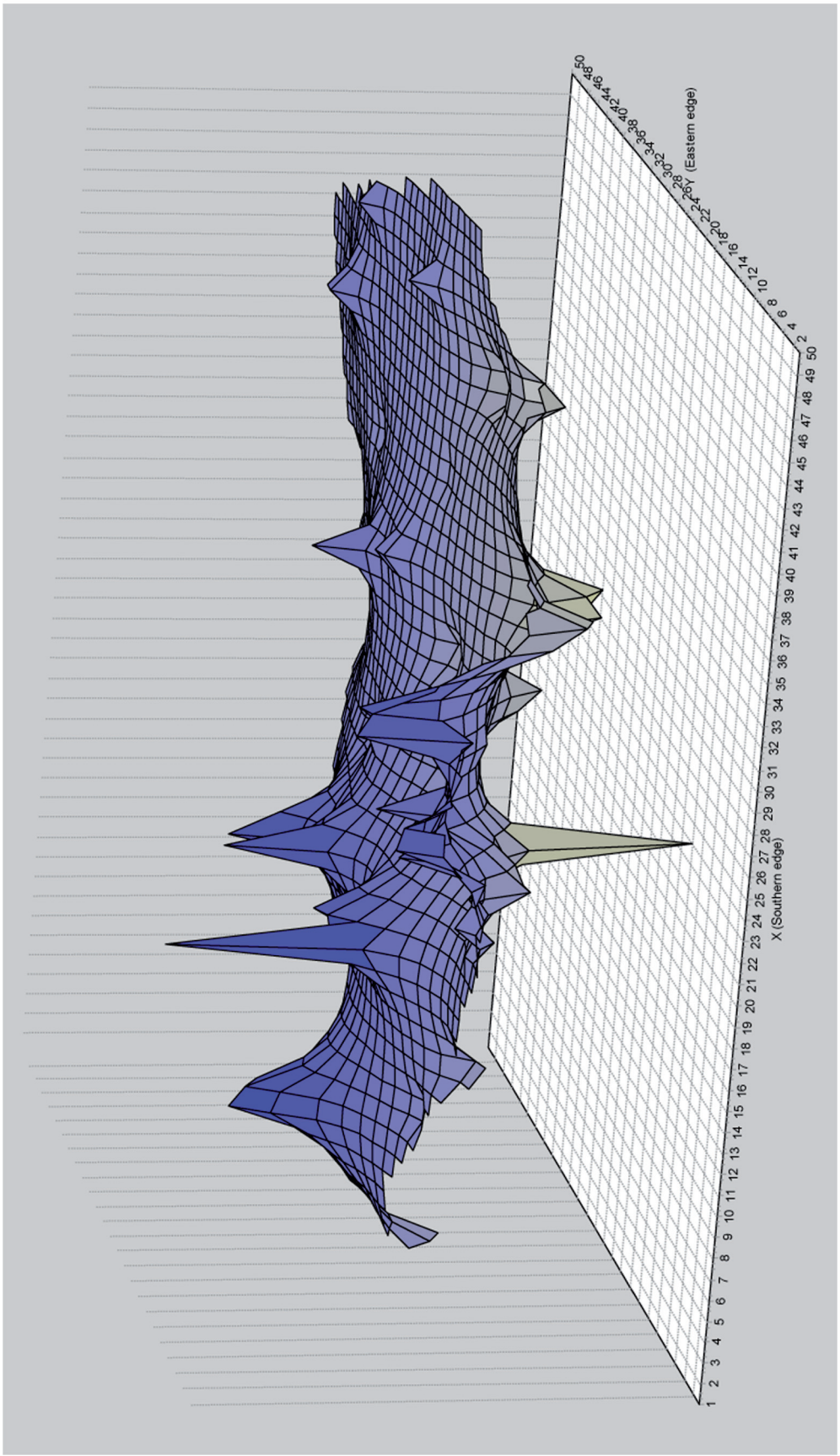

Fig. 5. Genetic landscape interpolation of *M. persicae*. X, Y coordinates are not altitude/longtitude coordinates. Z coordinate is genetic distance.

geographic distance ($r = 0.6392$, $P = 0.0070$) (Cao et al. 2012). Andrew et al. (2009) reported that genetic isolation did not significantly correlate with geographic distance ($R^2 = 0.02$; $P = 0.24$) in North American populations of *Aphis glycines* Matsumura. We speculated the possibility that geographic isolation exists among *M. persicae* populations in China, but isolation by distance could not fully explain our results. Rivers and mountains in China may, to a certain extent, create an effect on the separation of different geographic populations that linear distance did not reflect. For the genetic landscape interpolation, the emergence of genetic distance peak means the appearance of geographic barriers. The big genetic distance variations between populations from the northwest to the southern may be related directly to the complex geological structure locally (such as plateaus, basins and mountainous). The stable genetic distance among east populations indicates extensive gene flow. Factors that influence gene flow, such as transportation, trade, and migration, were also not expressed by the linear distance. And geographic isolation may also correlate with specific aphid species.

Our study was the first to comprehensively analyze the population structure of *M. persicae* (Sulzer) in China using microsatellite markers. We found that the northern and southern populations had clear genetic differences, but the genetic relationships between pairwise populations were difficult to directly survey and could only be indirectly speculated by the genetic information. Degree to which factors affect population structures is difficult to measure accurately. In future studies, additional molecular markers will be adopted to obtain more accurate and reliable genetic information. Such as mitochondrial DNA markers, so that can make sure the exchange between populations and migratory situation.

Acknowledgments

We thank Linzhen Cao for technical help. This work was supported by the earmarked fund for Modern Agro-Industry Technology Research System grants CARS-29-08.

References Cited

Andrew, P. M., W. Zhang, K. J. Jung, T. S. Kang, and M.A.R. Mian. 2009. Population genetic structure of *Aphis glycines*. Environ. Entomol. 38: 1301–1311.

Blackman, R. L., and F. V. Eastop. 1984. Aphids on the World's crops: an identification and information guide. John Wiley and Sons, Chichester.

Blackman, R. L., and F. V. Eastop. 2000. Aphids on the world's crops: an identification and information guide. Wiley, Chichester, UK.

Bohonak, A. J. 2002. IBD (isolation by distance): a program for analyses of isolation by distance. J. Hered. 93: 153–154.

Bostein, D., L. R. Whlte, M. Skolniek, and W. R. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32: 314–331.

Cao, J. J., J. Li, Q. J. Niu, X. X. Liu, and W. Q. Zhang. 2012. Population structure of *Aphis spiraecola* (Hemiptera: Aphididae) on pear trees in China identified using microsatellites. J. Econ. Entomol. 105: 583–591.

Delmotte, F., N. Leterme, P. J. Gauthier, C. Rispe, and C. J. Simon. 2002. Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. Mol. Ecol. 11: 711–723.

Doyle, J. J., and L. J. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. 1: 47–50.

Felsenstein, J. 2004. PHYLIP (Phylogeny Inference Package) version 3.6. Department of Genome Sciences, University of Washington, Seattle, WA. (http://evolution. genetics. washington. edu/phylip/phylipweb. html).

Fenton, B., J.A.T. Woodford, and G. Malloch. 1998. Analysis of clonal diversity of the peach-potato aphid, *Myzus persicae* (Sulzer), in Scotland, UK and evidence for the existence of a predominant clone. Mol. Ecol. 7: 1475–1487.

Fenton, B., G. Malloch, M. Navajas, J. Hillier, and A.N.E. Birch. 2003. Clonal composition of the peach-potato aphid *Myzus persicae* (Homoptera: Aphididae) in France and Scotland: comparative analysis with IGS fingerprinting and microsatellite markers. Ann. Appl. Biol. 142: 255–267.

Fuentes-Contreras, E., C. C. Figueroa, M. Reyes, M. L. Briones, and M. H. Niemeyer. 2004. Genetic diversity and insecticide resistance of *Myzus persicae* (Hemiptera: Aphididae) populations from tobacco in Chile: evidence for the existence of a single predominant clone. Bull. Entomol. Res. 94: 11–18.

Guillemaud, T., L. Mieuzet, and C. J. Simon. 2003. Spatial and temporal genetic variability in French populations of the peach–potato aphid, *Myzus persicae*. Heredity 91: 143–152.

Han, M. H., M. Y. Wang, Z. Wang, and R. Z. Shen. 2009. Genetic differentiation of *Myzus persicae* based on host plants revealed by using microsatellites. Chin. Bull. Entomol. 46: 244–249 (in Chinese).

Jarne, P., and P.J.L. Lagoda. 1996. Microsatellite, from molecules to populations and back. Trends Ecol. Evol. 11: 45–52.

John, T. M., K. Louise, L. M. Gaynor, and F. Brian. 2009. Tracking the global dispersal of a cosmopolitan insect pest, the peach potato aphid. BMC Ecol. 9: 1–13.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc. Natl. Acad. Sci. USA. 104: 3883–3888.

Liu, Y. G., H. Y. Qi, X. H. Li, K. M. Cuo, P. H. Lv, L. J. Zhang, and G. C. He. 2010. Genetic similarity of *Myzus persicae* from different geographical populations in Gansu province. Scientia Agricultura Sinica 43: 3134–3142 (in Chinese).

Llewellyn, K. S., D. H. Loxdale, R. Harrington, P. C. Brookes, J. S. Clark, and P. Sunnucks. 2003. Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. Mol. Ecol. 12: 21–34.

Maudet, C., C. Miller, B. Bassano, C. Breitenmoser-Würsten, D. Gauthier, G. Obexer-Ruff, J. Michallet, P. Taberlet, and G. Luikart. 2002. Microsatellite DNA and recent statistical methods in wild conservation management: application in Alpine ibex Capra ibex (ibex). Mol. Ecol. 11: 421–436.

Miller, M. P. 2005. Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. J. Hered. 96: 722–724.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106: 283–292.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Novak, S. J., and N. R. Mack. 2005. Genetic bottlenecks in alien plant species, pp. 201–228. *In* D.F. Sax, J.J. Stachowicz, and S.D. Gaines (eds.), Species invasions: insights into ecology, evolution, and biogeography. Sinauer Associates, Sunderland, MA.

Ordonez, M. E., and A. J. Kolmer. 2007. Simple sequence diversity of a worldwide collection of *Puccinia triticina* from durum wheat. Phytopathology 97: 574–583.

Papura, D., C. J. Simon, F. Halkett, F. Delmotte, and J.F. Le Gallic. 2003. Predominance of sexual reproduction in Romanian populations of the aphid *Sitobion avenae* inferred from phenotypic and genetic structure. Heredity 90: 397–404.

Peakall, R., and E. P. Smouse. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288–295.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Simon, J. C., S. Carre, M. Boutin, N. Prunier-Leterme, and B. Sabater-Mu-Nuoz. 2003. Host-based divergence in populations of the pea aphid: insights from nuclear markers and the prevalence of facultative symbionts. Proc. R. Soc. London B 270: 1703–1712.

Sloane, M. A., P. Sunnucks, A.C.C. Wilson, and F. D. Hales. 2001. Microsatellite isolation, linkage group identification and determination of recombination frequency in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). Genet. Res. 77: 251–260.

Sokal, R. R., and J. F. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. W. H. Freeman, New York.

Van den Heuvel, J.F.J.M., M. Verbeek, and F. van der Eilk. 1994. Endosymbiotic bacteria associated with circulative transmission of potato leafroll virus by *Myzus persicale*. J. Gen. Virol. 75: 2559–2565.

Van Emden, H. F., F. V. Eastop, D. R. Hughes, and J. M. Way. 1969. The ecology of *Myzus persicae*. Annu. Rev. Entomol. 14: 197–120.

Vorburger, C. 2006. Temporal dynamics of genotypic diversity reveal strong clonal selection in the aphid *Myzus persicae*. J. Evol. Biol. 19: 97–107.

Vorburger, C., M. Lancaster, and P. Sunnucks. 2003. Environmentally related patterns of reproductive modes in the aphid *Myzus persicae* and the predominance of two 'superclones' in Victoria, Australia. Mol. Ecol. 12: 3493–3504.

Vorwerk, S., and A. Forneck. 2006. Reproductive mode of grape phylloxera (*Daktulosphaira vitifoliae*, Homoptera: Phylloxeridae) in Europe: molecular evidence for predominantly asexual populations and a lack of gene flow between them. Genome 49: 678–688.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Wilson, A.C.C., P. Sunnucks, L. R. Blackman, and F. D. Hales. 2002. Microsatellite variation in cyclically parthenogenetic populations of *Myzus persicae* in south-eastern Australia. Heredity 88: 258–266.

Wilson, A.C.C., P. Sunnucks, and F. D. Hales. 2003. Heritable genetic variation and potential for adaptive evolution in asexual aphids (Aphidoidea). Biol. J. Linnean Soc. 79, 115–135.

Wright, S. 1978. Evolution and genetic populations, vol. 4. Variability within and among natural populations, pp. 79–103. University of Chicago Press, Chicago and London.

Yang, X. W., and X. X. Zhang. 1999. RAPD-PCR analysis of population differentiation of green peach aphid in China. Acta Entomologica Sinica 42: 372–380 (in Chinese).

Yeh, F. C., C. R. Yang, T.B.J. Boyle, H. Z. Ye, and X. J. Mao. 1999. POPGENE: The user-friendly shareware for population genetic analysis. University of Alberta, Edmonton, Canada.

Zhang, G. X., and S. T. Zhong. 1983. Economic insect fauna of China, vol. 25. Homoptera: Aphidoidae, pp. 312–313. Science Press, Beijing (in Chinese).

Received 9 September 2013; accepted 5 January 2014.