

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Ma, L;Cao, LJ;Hoffmann, AA;Gong, YJ;Chen, JC;Chen, HS;Wang, XB;Zeng, AP;Wei, SJ;Zhou, ZS

Title:

Rapid and strong population genetic differentiation and genomic signatures of climatic adaptation in an invasive mealybug

Date:

2020-05-01

Citation:

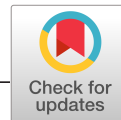
Ma, L., Cao, L. J., Hoffmann, A. A., Gong, Y. J., Chen, J. C., Chen, H. S., Wang, X. B., Zeng, A. P., Wei, S. J. & Zhou, Z. S. (2020). Rapid and strong population genetic differentiation and genomic signatures of climatic adaptation in an invasive mealybug. *Diversity and Distributions*, 26 (5), pp.610-622. <https://doi.org/10.1111/ddi.13053>.

Persistent Link:

<https://hdl.handle.net/11343/252354>

License:

[cc-by](#)



# Rapid and strong population genetic differentiation and genomic signatures of climatic adaptation in an invasive mealybug

Ling Ma<sup>1,2,3</sup> | Li-Jun Cao<sup>2</sup> | Ary A. Hoffmann<sup>4</sup> | Ya-Jun Gong<sup>2</sup> | Jin-Cui Chen<sup>2</sup> | Hong-Song Chen<sup>5</sup> | Xu-Bo Wang<sup>6</sup> | Ai-Ping Zeng<sup>3</sup> | Shu-Jun Wei<sup>2</sup> | Zhong-Shi Zhou<sup>1</sup>

<sup>1</sup>State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

<sup>2</sup>Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

<sup>3</sup>Institute of Insect Science, Hunan Agriculture University, Changsha, China

<sup>4</sup>School of BioSciences, Bio21 Institute, The University of Melbourne, Melbourne, Victoria, Australia

<sup>5</sup>Guangxi Key Laboratory for Biology of Crop Diseases and Insect Pests, Institute of Plant Protection, Guangxi Academy of Agricultural Sciences, Nanning, China

<sup>6</sup>Yunnan Academy of Biodiversity, Southwest Forestry University, Kunming, China

## Correspondence

Shu-Jun Wei, Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China.  
Email: shujun268@163.com

Zhong-Shi Zhou, State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China.  
Email: zs.zh@126.com

## Funding information

National Key Research and Development Program of China, Grant/Award Number: 2016YFC1202100; Natural Science Foundation of Beijing Municipality, Grant/Award Number: 6184037 and 6162010; Beijing Key Laboratory of Environmentally

## Abstract

**Aim:** A growing number of studies suggest that adaptation of invasive species plays key roles in their successful establishment in novel environments. However, adaptation of invasive species to climatic conditions remains poorly characterized. This study aimed to understand the population genetic structure produced by the cotton mealybug *Phenacoccus solenopsis* invasion and to identify preliminary signals of selection during its range expansion.

**Location:** China.

**Methods:** We examined genetic structure of 11 populations across China using SNPs, microsatellites and a segment of mitochondrial *cox1* gene. ADMIXTURE, STRUCTURE and DAPC were used to infer population genetic structure; the dispersal routes were reconstructed by the DIYABC; SNPs potentially related to climate adaptation were identified by using four populations differentiation methods and three environmental association methods.

**Results:** Strong genetic differentiation was found among populations with  $F_{ST}$  values ranging from 0.097 to 0.640 based on SNPs. Populations located at the northern expansion edge exhibited the highest genetic differentiation and the lowest genetic diversity. Demographic analyses indicated that all populations were introduced from a single source population with small effective size and low recent gene flow. RDA analysis showed that climatic variables explained a higher proportion of genetic variance (43%) compared to population structure variables (15%). The top climatic variables associated with genetic differentiation were precipitation of the mean temperature of warmest quarter, mean temperature of driest quarter and isothermality. Genes related to climate candidate SNPs were mainly enriched to pathways of development, energy and xenobiotic metabolisms.

**Main conclusions:** We found that extremely rapid and strong population genetic differentiation among populations appears to have developed after introduction in the

L. Ma and L.-J. Cao authors contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd

Friendly Pest Management on Northern Fruits, Grant/Award Number: BZ0432; BAAFS-UOM Joint Laboratory on Pest Control Research

Editor: Zhonghe Hou

cotton mealybug. Our study points to rapid neutral evolution and suggests possible climatic adaptation despite low genetic diversity in this invasive species.

#### KEYWORDS

adaptation, insects, invasive species, phylogeography, population genetics

## 1 | INTRODUCTION

Invasive species commonly show a high potential to spread across vast geographical areas with diverse environmental conditions (Seebens et al., 2017). However, invasive populations dispersed to new environments often face strong novel selective pressures (Liebhold & Tobin, 2008; Renault, Laparie, Mccauley, & Bonte, 2018). Previous studies showed that these species usually exhibit higher phenotypic plasticity than native species (Amy Michelle, Michael, & Nicotra, 2011; Felden et al., 2018; Vanwallendael et al., 2018), helping them to colonize new areas. In addition, both theoretical and empirical studies suggest that invasive populations of many species are expected to undergo evolutionary adaptation to respond to new conditions (Andrew, Jensen, Hagen, Lundregan, & Griffith, 2018; Lee, 2002; Renault et al., 2018; van Boheemen & Atwater, 2019; Willoughby et al., 2018). While this evolutionary potential may be curtailed as a consequence of bottlenecks limiting genetic diversity, most invasive populations are still expected to have the potential to adapt rapidly through evolution (Bock et al., 2015; Estoup et al., 2016). This means that invasive populations provide an opportunity to understand and contrast population evolution under both neutral (bottlenecks, gene flow) and adaptive processes (Bock et al., 2015; Lee, 2002).

Among many aspects of the environment that can restrict the distribution and abundance of invasive species, climatic conditions are likely to be particularly important, posing strong selective pressures on invasive populations (Chown et al., 2015; Colautti & Barrett, 2013; Renault et al., 2018; Roura-Pascual et al., 2011; van Boheemen & Atwater, 2019). Invertebrate species are particularly strongly influenced by climatic conditions on many biological characters, such as reproduction, developmental durations, diapause and survival. Studies have shown differences in climatic adaptation between invasive and non-invasive populations of invertebrates (Goubert et al., 2017; Hill, 2013; Jarošík, Kenis, Honěk, Skuhrovec, & Pyšek, 2015; Urbanski et al., 2012) and comparisons of species distributions in native and invaded regions provide additional indirect support for adaptive evolutionary shifts (Hoffmann, 2017). In newly colonized areas, climatic differences between early invaded areas and distribution edges may lead to differed selective pressures, and the opportunity to investigate rates of climatic adaptation in invading invertebrates.

The cotton mealybug, *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae), is a damaging and emerging invasive pest with harmful effects on agricultural and ornamental plants, with the pest attacking approximately 60 families of plants, including crop species such as cotton (Ibrahim, Moharum, & Elghany, 2015; Wang et al.,

2010). Originating from North America, this species has spread rapidly to 24 countries. Its distribution was restricted in cold areas as shown by its native range, and its potential distribution predicted from climatic models (Wang et al., 2010). In China, it was first reported in Guangdong province in 2008 (Lu, Zeng, Wang, Xu, & Chen, 2008; Wu & Zhang, 2009). Within ten years, it had rapidly spread to many areas of China and reached its predicted distribution range, posing a severe threat to cotton production (Wei et al., 2017).

All developmental stages of *P. solenopsis* are wingless, and most times, they are attached to the host plant, except for a short-lived flying male stage (Lu et al., 2008), limiting the distance covered by active dispersal. Long-distance dispersal of this species is mainly mediated by humans transporting infested plants, and thus, ongoing gene flow may be among populations. The distribution of *P. solenopsis* is mainly restricted by temperature and humidity (Dhawan, Singh, Aneja, & Saini, 2009; Wang et al., 2010). In areas with dry conditions, *P. solenopsis* occurs more commonly on roots, stems and foliage close to the soil line, while in humid areas, it settles on the upper foliage of the plant (Hodgson, Abbas, Arif, Saeed, & Karar, 2008). In China, *P. solenopsis* mainly invaded into southern areas with relatively high temperatures and humidity. However, it also established a population in 2011 in Xinjiang of northwestern China, where there are low temperatures and dry climatic conditions, posing potentially strong selective pressures on the population.

In this study, we examined population genetic structure and genomic signals of climatic selection of *P. solenopsis* across populations from China using SNPs, microsatellites and mitochondrial DNA. Our aim was to understand the population genetic structure produced by the *P. solenopsis* invasion and to identify preliminary signals of selection. We tested whether the low mobility of *P. solenopsis* impacted genetic structure and whether there were genomic signatures of adaptation under strong climatic pressure, especially in populations of the northern expansion of this species.

## 2 | METHODS

### 2.1 | Sample collection and DNA extraction

Female adults of *P. solenopsis* were collected from 11 locations across its distribution areas in China in 2017 (Table S1, Figure 1a). Genomic DNA was extracted from the whole individual of *P. solenopsis* using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) with an additional step of RNase treatment. Voucher DNA of specimens was stored in the Integrated Pest Management Laboratory of Beijing Academy of Agriculture and Forestry Sciences at  $-80^{\circ}\text{C}$ .

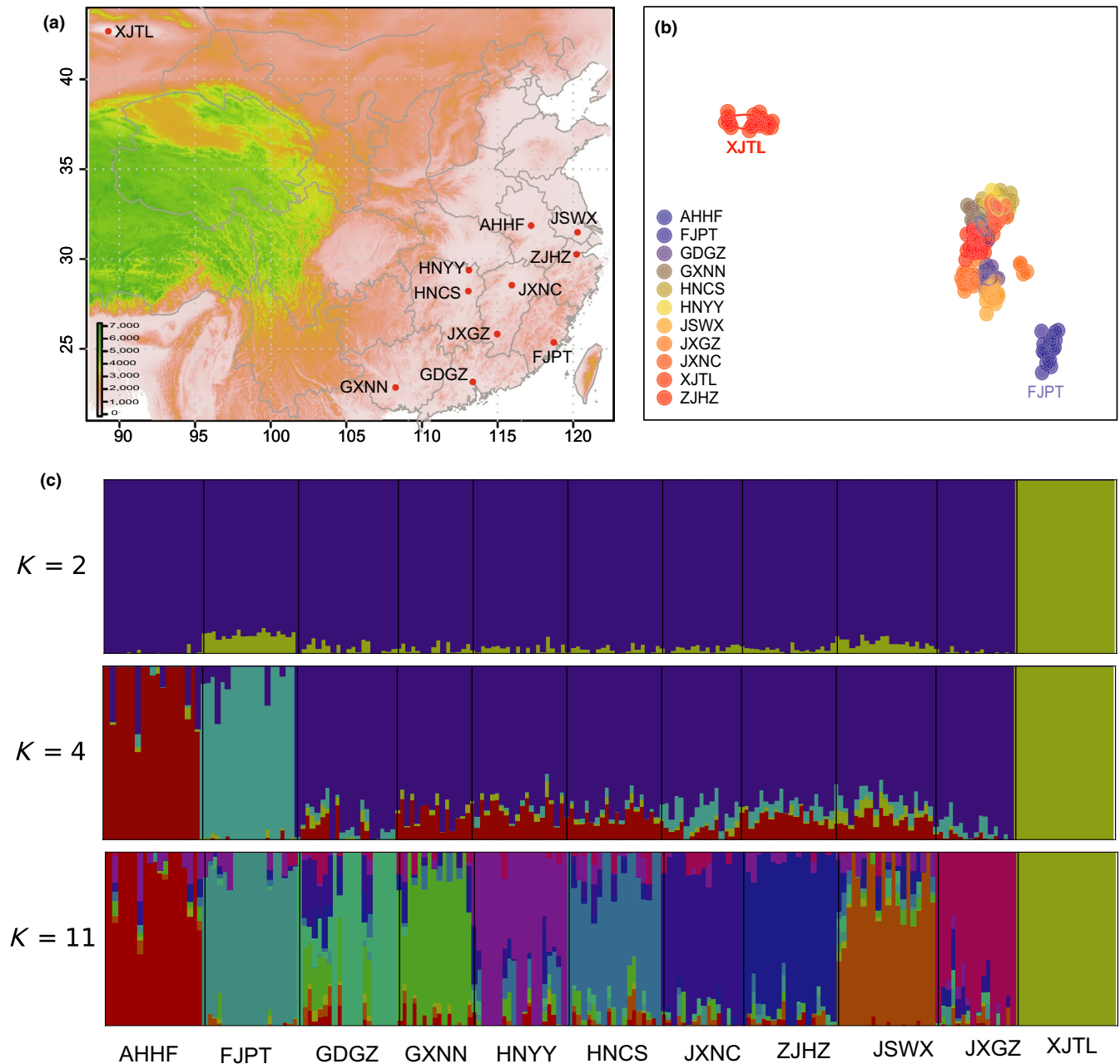


**FIGURE 1** Sampling locations (a) and genetic structure of *Phenacoccus solenopsis* populations (b, c) based on potentially neutral genetic SNPs. JXNC: Nanchang, Jiangxi; JXGZ: Ganzhou, Jiangxi; FJPT: Putian, Fujian; ZJHZ: Zhejiang, Hangzhou; HNY: Yueyang, Hunan; HNCS: Changsha, Hunan; XJTL: Tulufan, Xinjiang; AHHF: Hefei, Anhui; JSWX: Wuxi, Jiangsu; GDGZ: Guangdong, Guangzhou; GXNN: Nanning, Guangxi. The DAPC analysis showed XJTL and FJPT are outlier populations (b); The ADMIXTURE analysis showed XJTL, FJPT and AHHF were separated into different clusters when K was increased from 2 to 4, while when K was increased to 11, almost all populations were assigned into different clusters except for HNCS and JXNC (c) [Correction added on 30 May 2020, after first online publication: the location of red dots inside Figure 1a were previously incorrect and have been updated in this version.]

## 2.2 | SNP genotyping

We used the double-digest restriction site-associated DNA (ddRAD) method to develop genome-wide SNP markers (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). To select the endonuclease enzymes for digestion, we simulated digestion using the draft genome of *P. solenopsis* (Ma et al., 2019) as a reference through the *DDSilico* program (Rašić, Filipović, Weeks, & Hoffmann, 2014).

Libraries were constructed for each individual as described elsewhere (Peterson et al., 2012). Twenty individuals from each population were used for genotyping. In brief, 120 ng of genomic DNA in a 50 µl volume was digested with *Nla*III and *Ac*II restriction enzymes (New England Biolabs) for 3 hr at 37°C. Restricted DNA was purified with 75 µl (1.5×) SpeedBeads (GE), and each individual was ligated to a unique pair of modified Illumina P1 (5 bp) and P2 adapters (4 bp) overnight at 16°C, followed by a heat-deactivation

step under the conditions of 65°C for 10 min, with 22 cycles at 20°C of 1 min. These products with unique adapter of one population were pooled and cleaned with beads. We then selected cleaned DNA fragments with a size of 420–540 bp by BluePippin on a 2% gel cassette (Sage Sciences). The selected pooled DNA was amplified in 20 µl reactions containing 5 µl DNA, 10 µl of PCR Phusion 2× master mix, 0.8 µl of forward primer (50 µM) and reverse primer (50 µM) as well as 3.4 µl H<sub>2</sub>O using the conditions: 30 s at 98°C, followed by 12 cycles at 98°C for 10 s, 65°C for 30 s, 72°C for 45 s, a post-cycle incubation at 72°C for 5 min and holding at 4°C. Final library was purified by 16 µl 0.8× SpeedBeads. Quality of ddRAD libraries was evaluated by Qubit 3.0 and Agilent Bioanalyses 2100, and sequenced by Illumina HiSeq 4000 platform to generate 150-bp paired-end reads.

We used STACKS version 2.0 for SNP calling (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). FASTQC version 1.1.5 (Andrews, 2013) was used to estimate the content of GC and quality of raw sequences. Raw reads with sufficiently high sequencing quality and correct barcode were retained. After quality control, ddRAD barcodes were removed by the process\_radtags program (Catchen et al., 2013). Sequences were aligned to draft genome of *P. solenopsis* using BOWTIE version 2.2.9 (Langmead & Salzberg, 2012) with default parameters and exported the highest aligned core of SAM files. After removal of lower mapping rate individuals (<80%), 210 samples were kept and processed for SNP identification by the *ref\_map.pl* program (Catchen et al., 2013) with a minimum depth of five and a maximum likelihood framework.

SNPs meeting the following criteria were exported using *populations*: (a) SNPs that were included in at least 80% samples of a population; (b) SNPs with a minor allele frequency (MAF) higher than 0.05; and (c) SNPs with observed heterozygosity >0.8. The final data sets were filtered using R package *vcfR* (Knaus & Grünwald, 2017) by retaining SNPs with coverage at least 10 X, a missing rate in one sample <5%, and samples with a missing rate of SNPs <20%.

### 2.3 | Microsatellite genotyping and mitochondrial DNA sequencing

As a complement to SNPs, we genotyped 21 microsatellite loci from 24 individuals of *P. solenopsis* per population following the method described in Ma et al. (2019) (Table S2). Briefly, a universal PC tail (5'CAGGACCAGGCTACCGTG3') (Blacket, Robin, Good, Lee, & Miller, 2012) was used to identify forward primer candidates for amplification. All PCR products were analysed using an ABI 3730xl DNA Analyzer (Applied Biosystems) with the GeneScan 500 LIZ size standard (Applied Biosystems). Microsatellites were genotyped in GENEMAPPER version 4.0 (Applied Biosystems). We sequenced a segment of the mitochondrial *cox1* gene to validate the morphological identification of the specimens and level of genetic diversity on the mitochondrial genome, using the same primers, sequencing methods and analysis methods as mentioned previously (Ma et al., 2019).

### 2.4 | Outlier and neutral SNP identification

We used four population differentiation (PD) methods to identify outliers from all SNPs. First, BAYESCAN version 2.1 (Foll & Gaggiotti, 2008) was used to directly estimated the probability that a locus-specific  $F_{ST}$  was different to a population-specific  $F_{ST}$ . We ran this test with a prior odds value of 10, with 20 pilot runs of 5,000 iterations, a burn-in of 50 000 and with a thinning interval of 10. Loci with a false discovery rate (FDR) with <0.05 ( $q < 0.05$ ) were candidate outliers. Second, ARLEQUIN version 3.5.2.2 software (Excoffier & Lischer, 2010) was used to screen outliers with a *P* value lower than .01. Third, a Bayesian approach with the BayEnv model implemented in BAYPASS version 2.1 (Gautier, 2015) was used with default parameters. The XTX statistics were calibrated as suggested by Gautier (2015) and kept SNPs with values greater than a 99% threshold. Fourth, PCAdapt was used (Lotterhos & Whitlock, 2014), which has powerful and low positive rates in independent evaluations of various outliers methods. We ran PCAdapt with a burn-in of 200 steps,  $K = 10$ , and a false discovery rate (FDR) lower than 0.01.

The 19 bioclimatic variables at resolution 10 min of degrees were downloaded from the WorldClim database (<https://www.worldclim.org/>) for correlative analysis. To avoid high collinearity, we removed correlated predictors with  $|r| > .8$ . Six climatic variables remained for subsequent analysis, including bio3 (Isothermality), bio8 (Mean Temperature of Wettest Quarter), bio9 (Mean temperature of driest quarter), bio10 (Mean Temperature of Warmest Quarter), bio15 (Precipitation seasonality) and bio17 (Precipitation of Driest Quarter) (Figure S1).

We used three environmental association (EA) methods to identify outliers. First, BAYPASS version 2.1 was used for genotype–environment association analysis under the Basic Model. Second, latent factor mixed models (LFMM) implemented in the LEA R package (Eric, Schoville, Guillaume, & Olivier, 2013) were used to select outliers with  $|z|$ -scores and converted them into adjusted *p* values based on the Fisher–Stouffer method. Using the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995), we expected FDR correction at 5% to export candidate SNPs with adjusted *p* values <.05.

Third, we used a multivariate approach in a redundancy analysis (RDA) to estimate the extent to which the explained variance in SNP genotypes was explained by climatic variables and population structure and by their collinear portion (spatially autocorrelated climatic variation). RDA is a constrained linear ordination method that combines multiple linear regression and PCA (principal component analysis). Population structure effects were approximated by the first two principal components of a new PCA performed in R package. Two PCA vectors (PCA1 and PCA2) were remained when correlated predictors with  $|r| > .8$  were removed. Since spatial autocorrelation is usually reflected by neutral genetic structure, we did not include spatial covariates to avoid over-conditioning the model (Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). Both full RDA (population structure and climate) and partial redundancy analysis (pRDA) (population structure or climate) models were analysed in the R package *vegan*. The independent effect

of the environment was the variance values for the constrained matrix of population structure in the pRDA, while the independent effect of population structure was the equivalent for the constrained matrix of climate. The collinear proportion was calculated by subtracting the independent effects of environment and population structure from the total amount of variance explained in the full RDA model. We used the loadings of the SNPs in the ordination space to determine which SNPs are candidates for local adaptation (Forester, Lasky, Wagner, & Urban, 2018). Histograms of the loadings on each RDA axis usually show normal distributions. SNPs loading in the tails are more likely to be under selection as a function of predictors. So, we identified outliers as SNPs that load in the tails of these distributions with three times of standard deviation cut-off (two-tailed  $p = .0027$ ). We identified SNPs for local adaptation loading for the first three significantly constrained axes.

Potential PD SNPs were those identified in each of the four PD methods, and potential EA SNPs were those identified in each of the three EA methods. The candidate SNPs under selection were those identified by both EA and PD methods, while potential neutral SNPs were those not identified in any of the outlier analysis.

## 2.5 | Genetic diversity analysis

For SNPs, genetic diversity was measured using the *populations* program in STACKS version 2.0 (Catchen et al., 2013) and ARLEQUIN version 3.5.2.2 (Excoffier & Lischer, 2010), including nucleotide diversity ( $\pi$ ), average expected heterozygosity ( $H_e$ ), average observed heterozygosity ( $H_o$ ), and private alleles, and Wright's inbreeding coefficient ( $F_{IS}$ ).

For microsatellites, the genotyped data were corrected by MICRO-CHECKER (Van Oosterhout et al., 2004). The  $H_o$  and  $H_e$  were calculated by Microsatellite Tools (Park, 2001). Deviation from Hardy-Weinberg equilibrium (HWE), pairwise mean population differentiation ( $F_{ST}$ ) and inbreeding coefficients ( $F_{IS}$ ) were estimated in GENEPOP version 4.0.11 (Raymond & Rousset, 1995). We used FSTAT version 2.9.3 to test average of allelic richness ( $A_R$ ) of each locus (Goudet, 2017).

## 2.6 | Population genetic structure analysis

We estimated the individual ancestries using a maximum likelihood method implemented in ADMIXTURE version 1.3.0 (Alexander, Novembre, & Lange, 2009). Because ADMIXTURE's model does not explicitly account for linkage disequilibrium (LD) between markers (Alexander et al., 2009), we used all potentially neutral SNPs for analysis. Based on microsatellites, we used a Bayesian method implemented in STRUCTURE version 2.3.4 (Earl & vonHoldt, 2012) with a  $K$  from 2 to 11, replications of 30 times for each  $K$ , 200,000 Markov chain Monte Carlo iterations (MCMC) and a burn-in of 100,000 iterations. We inferred the number of clusters among populations

using discriminant analysis of principal component (DAPC) methods implemented in ADEGENET version 2.0.1 (Jombart, 2008) based on all SNPs and microsatellites, respectively. This method does not rely on any prior knowledge, while providing complementary results to assumption-based analyses of ADMIXTURE and STRUCTURE.

## 2.7 | Demographical history analysis

Contemporary effective population size for each population was estimated by NEESTIMATOR version 2.0 (Do et al., 2014) with LD model based on SNPs and microsatellites, respectively. Extended Bayesian Skyline Plot (EBSP) (Heled & Drummond, 2008; Trucchi et al., 2014) analysis was performed with BEAST2, which investigated effective population size changes through time using SNP data sets. Bottlenecks for each population were tested using BOTTLENECK version 1.2.02 (Piry, Luikart, & Cornuet, 1999) based on microsatellites with a two-phase model (TPM) and a stepwise mutation model (SMM).

We estimated recent and historical gene flow among populations using Bayesian methods implemented in BAYESASS version 3.0.4 (Han, Nalam, Yu, & Nachappa, 2019) and MIGRATE version 3.7.2 (Beerli & Felsenstein, 2001), respectively. BayesAss was used to estimate gene flow for the past few generations. Preliminary runs with 10,000,000 steps were performed to adjust mixing settings for allele frequencies and inbreeding coefficients. Then, ten runs of 100,000,000 steps with a burn-in of 50,000,000 were conducted. Three data sets with 400 randomly selected SNPs were used for analysis. Migrate was modelled to estimate asymmetric gene flow in all past times after the split of two populations. Mutation-scale effective population size ( $\theta = Ne\mu$ ) for each population and mutation-scale migration rate ( $M = m/\mu$ ) among all populations, where  $\mu$  is the mutation rate of genetic markers per generation, were simultaneously estimated with the Bayesian search strategy. Parameter values were as follows: long-chains = 1, long-inc = 20 (long sampling increment), long-sample = 100,000, burn-in = 100,000, heating = YES:1:(1.0:1.5:3.0:10,000.0), heated-swap = YES and replicate = YES:4. In the first run,  $\theta$  and  $M$  were estimated from  $F_{ST}$  values, while in subsequent runs, Bayesian estimates of  $\theta$  and  $M$  from the previous run were used.

Approximate Bayesian computation (ABC) analyses implemented in DIYABC version 2.1.0 (Cornuet et al., 2014) were used to estimate different scenarios on the introduction and dispersal routes of *P. solenopsis* in China using microsatellites. We performed independent analyses on ten pairs of population, whereas population XJTL was always included as suggested by Lombaert et al. (2014). Based on the results of genetic diversity and genetic structure, populations were divided into a north-western population (XJTL) and a southern population group. Three biologically scenarios were tested: (a) one was dispersed from the other population (scenario 1 and 2); (b) two populations were sequentially introduced from a source population (scenario 3); (c) one was introduced from a source population, and the other was independently derived from another source population (scenario 4) (Figure S2). Methods for

DIYABC analysis are described elsewhere (Wei et al., 2015), and priors for the parameters are provided in Table S3.

## 2.8 | Gene annotation of environment-associated outliers

Genes located 1 kb upstream and downstream of the candidate outliers associated with climate variables were extracted. Functions of the genes were annotated through orthology assignment by EGGNOG-MAPPER version 1.0.3 (Jaime et al., 2017); pathways were enriched by KOBAS version 3.0 (Chen et al., 2011).

## 3 | RESULTS

### 3.1 | Genotyping and genetic diversity

An average sequencing depth of 37.18× was obtained after data filtering and trimming (Table S4). The mapping rate ranged from 65.26% to 99.8% for each individual, with an average of 95.07%. We obtained 19,634 SNPs with a minimum coverage of 10×. There were 148 SNPs outliers identified by BayeScan, 1,020 by Arlequin, 1008 by BayPass, 193 by PCAdapt and in total 2,068 by four PD methods, while 497 SNPs were identified by three EA methods with 335 by BayPass, 83 by LFMM and 111 by RDA analysis. In total, 190 candidate EA SNPs were identified by both EA and PD analyses (Figure S3). Based on outlier scanning results, 7,277 SNPs were presumed as potentially neutral, while 2,375 SNPs were presumed under potential selection. Based on the estimated genome size of 261.61 Mb (Ma et al., 2019), the density of neutral SNPs was estimated to be 35.95 Kb/SNP.

There was a significantly higher  $H_o$  than  $H_e$  in eight of 11 population based on SNPs ( $p$  value of HWE test was lower than .05 and  $F_{IS}$  was negative, see Table 1). The nucleotide diversity ( $\pi$ ) ranged from 0.09 to 0.28. Two populations of XJTL and AHHF at the edge of the northern expansion had the lowest genetic diversity, and XJTL had the highest number of private alleles (Table 1).

For microsatellites, we found 46 alleles from 260 individuals. There was no significant departure from HWE and linkage disequilibrium for a locus in all populations or a population on all loci, except for deviations from HWE in FJTP and GDGZ based on all microsatellites ( $p < .01$ ) (Table S5). Negative  $F_{IS}$  (except for ZJHZ) and level of allele richness was found among all populations based on microsatellites (Table S5).

For the mitochondrial gene, all individuals shared the same haplotype as the one previously reported in Ma et al. (2019).

### 3.2 | Population differentiation and genetic structure

Pairwise  $F_{ST}$  values ranged from 0.097 to 0.640 and from 0.069 to 0.567 based on SNPs and microsatellites, respectively. The XJTL,

AHHF and FJPT populations showed the highest  $F_{ST}$  values with other populations (Table 2).

Population genetic structure analysis using ADMIXTURE showed that the 11 populations were optimally divided into 11 clusters. When  $K$  was increased from 3 to 11, XJTL was the first separated cluster, followed by AHHF and FJPT (Figure 1c). In DAPC analysis, results showed that XJTL and FJPT were clustered into two separate groups; other populations were assembled into one cluster (Figure 1b). STRUCTURE and DAPC analyses based on microsatellite loci generated similar results to those based on SNPs, indicating strong population differentiation of *P. solenopsis* and a high level of differentiation of the distantly distributed northwestern population of XJTL (Figure S4).

### 3.3 | Demographic history

All populations had a small effective population size. Population JXNC had the highest effective population size based on SNPs (38.5) and microsatellites (31.5). The other populations had an effective population size ranging from 2.4 to 34.4 based on SNPs, and 1.2 to 8.4 based on microsatellites (Table 1). Evidence for a bottleneck was found in the GDGZ, JXNC, FJPT and XJTL populations ( $p < .05$ ) (Table 1).

There was higher estimated historical gene flow from most southern populations (such as GDGZ, GXNN and JXNC), but low gene flow from the populations of XJTL, FJTP, ZJHZ and JSWX (Figure 2a). A low level of recent gene flow was found among most populations, except those from GXNN and HNYU to their neighbouring populations of HNYU, JXGZ and JXNC (Figure 2b, Table S6).

DIYABC analysis revealed that scenario 2 was supported with the highest posterior probability in 9 of 10 analyses, suggesting that the XJTL population was derived from southern populations in China. The posterior probabilities ranged from 0.6614 to 0.8755 (Table S7).

### 3.4 | Climatic effect on genetic variance

The full RDA model supported the role of climate and/or population structure in shaping the distribution of SNP genotypes ( $p = .001$ ;  $R^2 = .367$ ). In the partial RDA model, both climate conditioned on population structure ( $p = .001$ ;  $R^2 = .152$ ) and population structure conditioned on bioclimate ( $p = .001$ ;  $R^2 = .058$ ) showed significant effects on genetic variation. The first four RDA axes in the partial RDA model (where the climate was conditioned on population structure) were significant and together represented 79.61% of the constrained variance in SNP genotypes (Table S8). Climate accounted for 43.23% of the total explained genetic variance, population structure accounted for 15.90% and the collinear contribution of climate and population structure accounted for 40.87% (Table S9). When climate effects were conditioned by population structure effect in partial RDA analysis, the bio10 (mean temperature of warmest quarter) was highly correlated with genetic variance when the first two RDA axes were considered (Figure 3a), and

**TABLE 1** Genetic diversity, effective population size ( $N_e$ ) and  $p$  value of bottleneck test of *Phenacoccus solenopsis*

| Population | $H_o$ | $H_e$ | Pi   | Pri | HWE        | $F_{IS}$ | $N_e$ (95% confidence intervals) |                  | Bottleneck test  |                  |     |
|------------|-------|-------|------|-----|------------|----------|----------------------------------|------------------|------------------|------------------|-----|
|            |       |       |      |     |            |          | SNP                              | Microsatellite   | TPM <sup>a</sup> | TPM <sup>b</sup> | SMM |
| AHHF       | 0.18  | 0.15  | 0.15 | 0   | $p < .001$ | -0.21    | 2.8 (2.8–2.8)                    | 2.8 (1.7–7.8)    | 0.7              | 0.5              | 0.8 |
| FJPT       | 0.28  | 0.22  | 0.23 | 2   | $p < .001$ | -0.29    | 4.3 (4.3–4.3)                    | 8.4 (3.3–17.4)   | 0.0              | 0.0              | 0.0 |
| GDGZ       | 0.27  | 0.24  | 0.25 | 0   | $p < .001$ | -0.14    | 1.5 (1.5–1.5)                    | 1.8 (1.3–2.4)    | 0.0              | 0.0              | 0.0 |
| GXNN       | 0.24  | 0.23  | 0.23 | 0   | $p > .050$ | -0.06    | 3.1 (3.1–3.1)                    | 1.8 (1.2–2.7)    | 0.1              | 0.1              | 0.2 |
| HNCS       | 0.29  | 0.25  | 0.26 | 0   | $p < .001$ | -0.18    | 8.8 (8.8–8.8)                    | 3.2 (2.3–6.9)    | 0.2              | 0.1              | 0.3 |
| HNY Y      | 0.26  | 0.24  | 0.24 | 4   | $p < .001$ | -0.08    | 2.4 (2.4–2.4)                    | 2.5 (1.7–4.8)    | 0.5              | 0.5              | 0.7 |
| JSWX       | 0.26  | 0.21  | 0.22 | 1   | $p < .001$ | -0.25    | 8.9 (8.9–8.9)                    | 4.6 (2.4–9.9)    | 0.3              | 0.2              | 0.5 |
| JXGZ       | 0.30  | 0.24  | 0.25 | 5   | $p > .050$ | -0.28    | 13.3 (13.3–13.3)                 | 4.2 (2.4–8.8)    | 0.0              | 0.0              | 0.0 |
| JXNC       | 0.28  | 0.27  | 0.28 | 0   | $p > .050$ | -0.04    | 38.5 (38.4–38.6)                 | 31.5 (3.5–339.7) | 0.1              | 0.1              | 0.2 |
| XJTL       | 0.12  | 0.09  | 0.09 | 64  | $p < .001$ | -0.39    | 8.8 (8.8–8.9)                    | 1.2 (0.6–2.4)    | 0.0              | 0.0              | 0.0 |
| ZJHZ       | 0.32  | 0.25  | 0.26 | 0   | $p < .001$ | -0.35    | 34.4 (34.3–34.5)                 | 7.5 (3.1–15)     | 0.1              | 0.1              | 0.1 |

Note: Genetic diversity parameters were estimated from SNPs.

Abbreviations:  $F_{IS}$ , inbreeding coefficient. Bottleneck was tested based on microsatellites;  $H_e$ , expected heterozygosity;  $H_o$ , observed heterozygosity; HWE, Hardy–Weinberg equilibrium; IAM, infinite allele model; Pi, number of nucleotide diversity; Pri, number of private substitutions; SMM, stepwise mutation model; TPM, two-phase mode.

<sup>a</sup>TPM, (80% of SMM, 20% of IAM).

<sup>b</sup>TPM, (70% of SMM, 30% of IAM).

**TABLE 2** Pairwise  $F_{ST}$  among 11 *Phenacoccus solenopsis* population calculated from all SNPs (lower triangle) and microsatellites (upper triangle)

| Population | AHHF  | FJPT  | GDGZ  | GXNN  | HNCS  | HNY Y | JSWX  | JXGZ  | JXNC  | XJTL  | ZJHZ  |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| AHHF       |       | 0.334 | 0.226 | 0.280 | 0.147 | 0.178 | 0.169 | 0.347 | 0.188 | 0.543 | 0.136 |
| FJPT       | 0.451 |       | 0.177 | 0.281 | 0.199 | 0.235 | 0.218 | 0.318 | 0.207 | 0.347 | 0.256 |
| GDGZ       | 0.329 | 0.295 |       | 0.179 | 0.069 | 0.169 | 0.130 | 0.227 | 0.105 | 0.445 | 0.156 |
| GXNN       | 0.312 | 0.326 | 0.148 |       | 0.169 | 0.117 | 0.184 | 0.358 | 0.247 | 0.371 | 0.188 |
| HNCS       | 0.276 | 0.296 | 0.128 | 0.130 |       | 0.126 | 0.086 | 0.193 | 0.069 | 0.436 | 0.117 |
| HNY Y      | 0.271 | 0.301 | 0.150 | 0.117 | 0.097 |       | 0.121 | 0.306 | 0.166 | 0.441 | 0.172 |
| JSWX       | 0.329 | 0.311 | 0.179 | 0.192 | 0.145 | 0.165 |       | 0.329 | 0.173 | 0.410 | 0.165 |
| JXGZ       | 0.364 | 0.322 | 0.195 | 0.206 | 0.177 | 0.179 | 0.238 |       | 0.085 | 0.567 | 0.215 |
| JXNC       | 0.309 | 0.263 | 0.135 | 0.150 | 0.098 | 0.121 | 0.167 | 0.153 |       | 0.483 | 0.119 |
| XJTL       | 0.640 | 0.557 | 0.490 | 0.521 | 0.488 | 0.502 | 0.520 | 0.537 | 0.490 |       | 0.452 |
| ZJHZ       | 0.277 | 0.277 | 0.132 | 0.122 | 0.111 | 0.109 | 0.151 | 0.165 | 0.119 | 0.482 |       |

Note: Pairwise  $F_{ST}$  was estimated by ARLEQUIN version 3.5 with 100 permutations. There is significant difference in pairwise  $F_{ST}$  values ( $p < .05$ ). Value of lower triangle was calculated from 19,634 SNPs, and value of upper triangle was calculated from microsatellites.

the bio3 (isothermality) was highly correlated with genetic variance when the RDA3 and RDA4 axes were considered (Figure 3b).

### 3.5 | Candidate genes related to climatic adaptation

There are 190 candidate SNPs obtained by both EA and PD methods. In total, 33 genes were obtained related to candidate loci, and 26 genes were annotated by eggNOG-mapper, which were enriched to energy metabolism (carboxylesterase and esterase), growth and development (ChBD, chitin-binding peritrophin-A domain), transcription factors

(bHLH, basic helix-loop-helix), detoxification metabolism (ABC transporter and UDP-glycosyltransferase activity) (Tables S10, S11). In RDA analysis, we obtained 111 SNPs associated with climate variables. The bio10 variable had the most associated loci, followed by bio9 (mean temperature of driest quarter) and bio3 (Figure S5).

## 4 | DISCUSSION

In this study, we used genome-wide SNPs, microsatellites and mitochondrial DNA to examine the invasion genetics of the

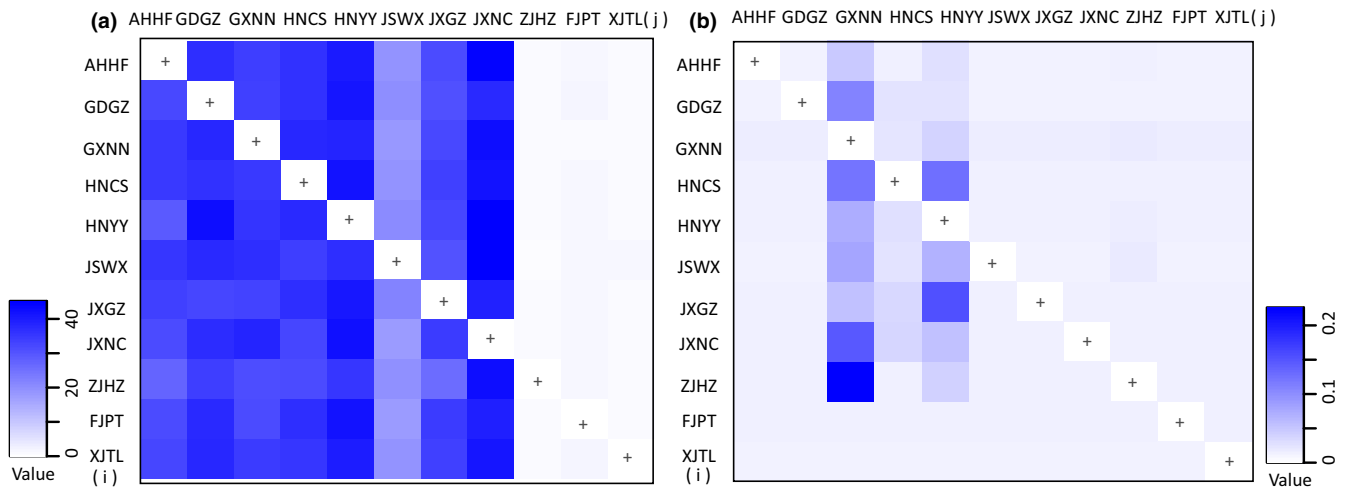


**FIGURE 2** Historical (a) and recent gene flow (b) among 11 populations of *Phenacoccus solenopsis* based on SNPs. The direction was from population  $j \rightarrow i$  (see axes labels). Dark colour indicates a high level of gene flow, while the light colour indicates a low level of gene flow. Historical gene flow was represented by effective migrants per generation, which equals to mutation-scale effective population size ( $\theta = Ne\mu$ ) multiplying mutation-scale migration rate ( $M = m/\mu$ ) calculated from Migrate. Recent gene flow was represented by migration rate ( $m$ ) calculated from BayesAss

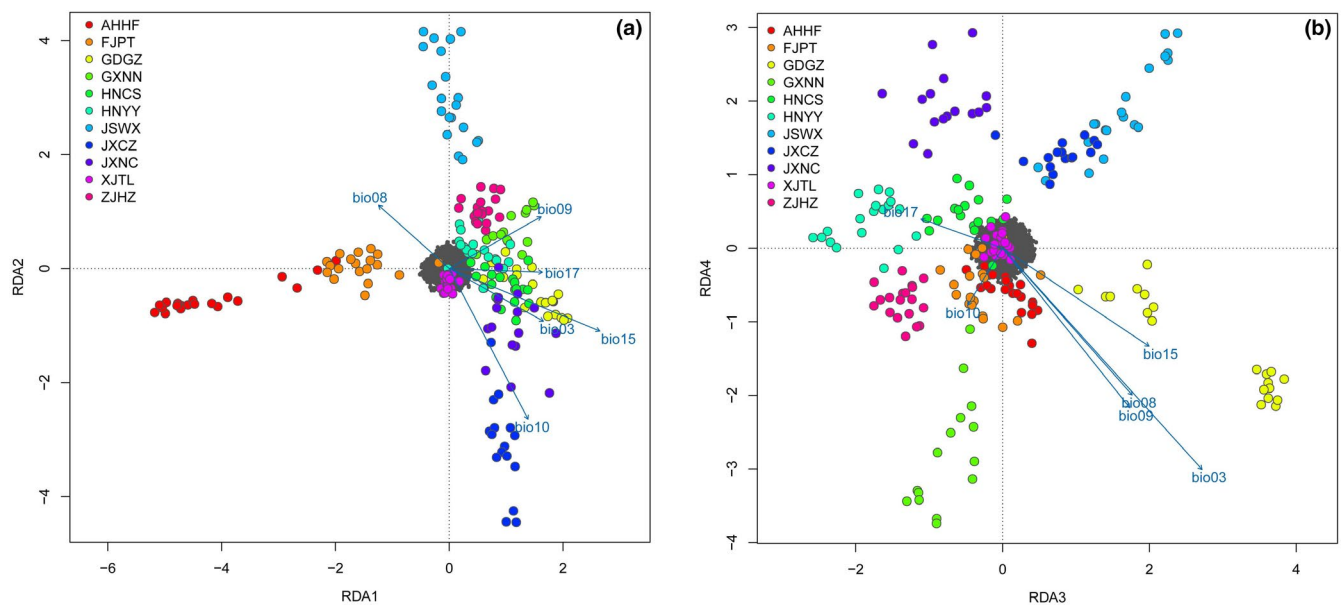


**FIGURE 3** Partial RDA analysis on genetic variance explained by the effects of climate when population structure effects were constrained. Individuals from the same population were indicated by circles with the same colour. The blue vectors are the environmental predictors. The relative arrangement of individuals and climate variables in the ordination space reflects their relationship with the ordination axes, which are linear combinations of the predictor variables. Combinations of the first four ordination axes were shown, that is RDA1 and RDA2 (a), RDA3 and RDA4 (b). Six climate variables were used, including bio3 (Isothermality), bio8 (Mean temperature of wettest quarter), bio9 (Mean temperature of driest quarter), bio10 (Mean temperature of warmest quarter), bio15 (Precipitation seasonality) and bio17 (Precipitation of driest quarter)

invasive species *P. solenopsis* in its recently introduced area of China. The ddRAD approach provides an efficient method to develop high-throughput SNPs, including potentially neutral and selected loci (Pujolar et al., 2013); the microsatellites provided a supplement to SNPs for making neutral process inferences, such as about population genetic structure and demographic history; the mitochondrial *cox1* gene helped to validate the morphological

identification of the species. Differences in patterns of genetic diversity were found based on the three sets of genetic markers, suggesting the importance of genetic markers for population genetic analysis. Our study provided inferences about invasion routes and demographic history, as well as genomic variation potentially related to climatic adaptation of the invasive species in a novel environment.

#### 4.1 | Invasion history of this invasive species in China

Based on historical records, *P. solenopsis* first invaded Guangdong in 2008 (Wu & Zhang, 2009) and then rapidly spread to Hainan (2009), Fujian (2010), Zhejiang (2010), Hunan (2010), Jiangxi (2010), Guangxi (2010), Yunnan (2010), Sichuan (2010) and Xinjiang (2011) (Table S1). Historical records showed that it was introduced into Xinjiang more recently. The distribution of the invasive species in southern China is now continuous, but the Xinjiang population is an outlier far from the early invasion area of this pest and with different (cold and dry) climatic conditions. Historical records could not determine whether the Xinjiang population was introduced from adjacent countries.

Our population genetic approach provides supplemental information on the dispersal history of this invasive species, as has been the case for other invasive species (Arnaud & Thomas, 2010; Cao, Wei, Hoffmann, Wen, & Chen, 2016). The DIYABC analysis suggests that the Xinjiang population of *P. solenopsis* was likely introduced from southern China. This is consistent with what is known about the biology of the pest. All developmental stages of the invasive species attach on host plants except for flying male adults which have a very short life span (Zhu et al., 2011), making long-distance active dispersal of this species unlikely. The spread of this species is mainly thought to be mediated by global integration, commercial trade and transportation of host plants (Ying et al., 2004). From a quarantine perspective, it had been detected on flowers *Hibiscus rosa-sinensis* and *Phalaenopsis aphrodite* in Xinjiang, with the plants originating from southern China. Although our study did not include populations from other countries, we explored this possibility by adding an unsampled ghost population in the DIYABC analysis (scenario 4). This approach can be an efficient method to account for incomplete sampling in population genetic studies (Cao et al., 2016; Eric et al., 2015). Nevertheless, testing samples of populations from adjacent countries will be required for confidence in this conclusion.

#### 4.2 | Rapid and strong genetic differentiation after the introduction

Although *P. solenopsis* has been in China for only about ten years, population genetic analysis based both on SNPs and microsatellites revealed strong population genetic differentiation, especially between the bridgehead populations involved in the northern expansion, such as XJTL and AHHF. Genome-wide SNPs here provided much higher resolution than microsatellites in detecting population structure, consistent with other comparisons of these marker systems (Rašić et al., 2014). Nevertheless, the level and pattern of population differentiation based on these marker systems was similar as expected if both are acting in a neutral manner at the population level. The highest  $F_{ST}$  values which were between XJTL and

other populations ranged from 0.45 to 0.64, while the lowest level of differentiation was 0.097 between two adjacent populations, estimated from SNPs (Table 2). Because all individuals used in our study shared one mitochondrial *cox1* haplotype, the strong genetic differentiation we observed is unlikely to be attributable to an unidentified cryptic species. This is an example of rapid and strong genetic differentiation developing in an invasive species.

Genetic differentiation of invasive species may be related to different source populations, bottlenecks, genetic drift and/or rapid adaptive evolution (Bertelsmeier & Keller, 2018; Cao et al., 2017; Dlugosch & Parker, 2010). Based on the demographic analysis, the population differentiation in the invasive species was mainly caused by evolutionary processes in China rather than involving different sources of introduced populations. The low mobility of most developmental stages on plants makes this species relatively easy to identify through quarantine screening, which may have limited the number of individuals dispersing into a new area, accounting for the small effective population size and limited recent gene flow. These factors can lead to rapid genetic differentiation among populations (Ellstrand & Elam, 1993).

#### 4.3 | Climatic adaptation

Genetic isolation by environmental conditions may generally be more common than isolation by geographical distance (Sexton, Hangartner, & Hoffmann, 2014). *Phenacoccus solenopsis* was discovered in 1898 by Tinsley in New Mexico, USA (Tinsley, 1898), and the potential distribution of this species appears to be limited by cold temperatures at high latitudes and altitudes, as well as aridity (Wang et al., 2010). The location of this species on the plant appears to be influenced by humidity (Hodgson et al., 2008). The current distribution of *P. solenopsis* in China has reached its predicted extent based on climatic models (Wang et al., 2010), but even so the cold and dry conditions experienced in northern China may impose strong stress on the invasive species, promoting evolutionary adaptation. Our multivariate analysis shows that variables related to both temperature and precipitation were associated with genetic patterns across populations of the invasive species, indicating the possible importance of environmental conditions overall. The genomic signatures provide possible indications on climate adaptation of *P. solenopsis*. Due to the confound influence of population structure, we were unable to exclude the neutral processes in multivariate analysis. Multiple sampling from different locations with the same climate conditions may help to reduce the influence of neutral variation in identification of adaptive loci.

In many species of mealybug, both sexes are capable of mating multiple times on the same day and on sequential days (Seabra et al., 2013). These mating behaviours may also be the case for *P. solenopsis* and help to reduce inbreeding and contribute to a higher effective population size. The *P. solenopsis* occurs multiple generations a year with development time from egg to adult stage about 25–30 days

(Prasad, Prabhakar, Sreedevi, Rao, & Venkateswarlu, 2012). These biological characters in the invasive species may increase the likelihood of rapid adaptation to a new environment (rapid generation time, increased genetic variability).

We found a higher observed heterozygosity than expected heterozygosity and negative  $F_{IS}$  in eight populations of *P. solenopsis* based on SNPs. Heterozygote excess can result from many potential causes, such as changes in effective population size, mating systems, asexual reproduction and overdominance (Stoeckel et al., 2006). Fewer breeders due to a small effective population size can result to heterozygote excess (Pudovkin, Zaykin, & Hedgecock, 1996). In *P. solenopsis*, there was a higher heterozygosity and smaller effective population size in XJTL with the opposite situation occurring in JXNC (Table 1), which may be consistent with this expected pattern. In male mealybugs, paternally derived chromosomes will be inactivated during spermatogenesis and therefore only the maternally derived euchromatic chromosomes are transmitted by males to the progeny, a phenomenon known as paternal genome elimination (PGE) (Seabra et al., 2013). PGE is a form of haplodiploid reproduction (Normark, 2004). Haplodiploid population is expected to be different from diploids in evolution, such as effective population size and stable polymorphisms (Hedrick & Parker, 1997). Whether this form of inheritance in mealybugs impacts on genetic diversity and rates of adaptive evolution requires further study.

To provide a preliminary test of potential adaptive patterns, we identified genes associated with climatic variables involved in metabolic processes and molecular processes. These functional genes identified in our study provide candidates for environmental adaptation of the invasive species which requires further work. For instance, it may be possible to compare the cold/aridity responses of this species and relate variation within and across populations to the SNP variation.

The invasion success of many species will depend on their ability to respond to natural selection (Colautti & Barrett, 2013; Lee, 2002). The recently developed method of characterizing genomic differences among populations provides an efficient approach to provide a preliminary test of local adaptation (Jensen, Foll, & Bernatchez, 2016; Munshi-South, Zolnik, & Harris, 2016; Pujolar et al., 2014) which can then be validated through additional work. However, few cases of adaptation of alien species to climatic conditions have been reported (Card et al., 2018). Our study provides SNP-based evidence of potential climatic adaptation in *P. solenopsis*, which may help to explain its invasiveness and its spatial expansion in recent years.

## 5 | CONCLUSION

Our study on the population genetics of an invasive species *P. solenopsis* revealed three major findings on the invasion of this species. First, the data are consistent with the invasive species invading China from a single source population. Second, extremely rapid and strong population genetic differentiation among populations appears to have developed after introduction, suggesting limited gene

flow, and small effective population sizes. Third, we found suggestive evidence of genomic signatures of climatic adaptation in an invasive species despite very low genetic diversity in the introduced population. Our study suggests rapid evolution in an invasive species which are worth investigating in additional studies with an experimental focus.

## ACKNOWLEDGEMENTS

The study was supported by the National Key Research and Development Program of China (2016YFC1202100), the Natural Science Foundation of Beijing Municipality (6184037, 6162010), the Beijing Key Laboratory of Environmentally Friendly Pest Management on Northern Fruits (BZ0432) and BAAFS-UOM Joint Laboratory on Pest Control Research.

## CONFLICT OF INTEREST

The authors declare they have no competing interest.

## DATA AVAILABILITY STATEMENT

The data and scripts supporting the results of this article are available in the Dryad repository (<https://doi.org/10.5061/dryad.0rxwdbrwf>).

## ORCID

Ary A. Hoffmann  <https://orcid.org/0000-0001-9497-7645>

Shu-Jun Wei  <https://orcid.org/0000-0001-7398-0968>

## REFERENCES

- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Amy Michelle, D., Michael, J., & Nicotra, A. B. (2011). Do invasive species show higher phenotypic plasticity than native species and if so, is it adaptive? A meta-analysis. *Ecology Letters*, 14, 419–431. <https://doi.org/10.1111/j.1461-0248.2011.01596.x>
- Andrew, S. C., Jensen, H., Hagen, I. J., Lundregan, S., & Griffith, S. C. (2018). Signatures of genetic adaptation to extremely varied Australian environments in introduced European house sparrows. *Molecular Ecology*, 27(22), 4542–4555. <https://doi.org/10.1111/mec.14897>
- Andrews, S. (2013). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics.
- Arnaud, E., & Thomas, G. (2010). Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology*, 19, 4113–4130. <https://doi.org/10.1111/j.1365-294X.2010.04773.x>
- Beerli, P., & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 4563–4568. <https://doi.org/10.1073/pnas.081068098>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57, 289–300.
- Bertelsmeier, C., & Keller, L. (2018). Bridgehead effects and role of adaptive evolution in invasive populations. *Trends in Ecology and Evolution*, 33, 527–534. <https://doi.org/10.1016/j.tree.2018.04.014>
- Blacket, M. J., Robin, C., Good, R. T., Lee, S. F., & Miller, A. D. (2012). Universal primers for fluorescent labelling of PCR fragments—an

- efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources*, 12, 456–463. <https://doi.org/10.1016/j.tree.2018.04.014>
- Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hubner, S., ... Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, 24, 2277–2297. <https://doi.org/10.1111/mec.13032>
- Cao, L. J., Wang, Z. H., Gong, Y. J., Zhu, L., Hoffmann, A. A., & Wei, S. J. (2017). Low genetic diversity but strong population structure reflects multiple introductions of western flower thrips (Thysanoptera: Thripidae) into China followed by human-mediated spread. *Evolutionary Applications*, 10, 391–401. <https://doi.org/10.1111/eva.12461>
- Cao, L. J., Wei, S. J., Hoffmann, A. A., Wen, J. B., & Chen, M. (2016). Rapid genetic structuring of populations of the invasive fall webworm in relation to spatial expansion and control campaigns. *Diversity and Distributions*, 22, 1276–1287. <https://doi.org/10.1111/eva.12461>
- Card, D. C., Perry, B. W., Adams, R. H., Schield, D. R., Young, A. S., Andrew, A. L., ... Castoe, T. A. (2018). Novel ecological and climatic conditions drive rapid adaptation in invasive Florida Burmese pythons. *Molecular Ecology*, 27, 4744–4757. <https://doi.org/10.1111/mec.14885>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chen, X., Xizeng, M., Jiaju, H., Yang, D., Jianmin, W., Shan, D., ... Liping, W. (2011). KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research*, 39, 316–322. <https://doi.org/10.1093/nar/gkr483>
- Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications*, 8, 23–46. <https://doi.org/10.1111/eva.12234>
- Colautti, R. I., & Barrett, S. C. (2013). Rapid adaptation to climate facilitates range expansion of an invasive plant. *Science*, 342, 364–366. <https://doi.org/10.1126/science.1242121>
- Cornuet, J. M., Pudlo, P., Veyssier, J., Dehnegarcia, A., Gautier, M., Leblois, R., ... Estoup, A. (2014). DIYABC v2.0: A software to make approximate bayesian computation inferences about population history using single nucleotide polymorphism. *DNA Sequence and Microsatellite Data. Bioinformatics*, 30, 1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>
- Dhawan, A. K., Singh, K., Aneja, A., & Saini, S. (2009). Distribution of mealy bug, *Phenacoccus solenopsis* Tinsley in cotton with relation to weather factors in South-Western districts of Punjab. *Journal of Entomological Research*, 1, 59–63.
- Dlugosch, K., & Parker, I. (2010). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17, 431–449. <https://doi.org/10.1111/j.1365-294X.2007.03538.x>
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Molecular Ecology Resources*, 14, 209–214.
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics*, 24, 217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>
- Eric, F., Schoville, S. D., Guillaume, B., & Olivier, F. O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30, 1687–1699. <https://doi.org/10.1093/molbev/mst063>
- Eric, L., Thomas, G., Jonathan, L., Robert, K., BenoT, F., Audrey, G., ... Emma, R. (2015). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, 23, 5979–5997.
- Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology and Systematics*, 47, 51–72. <https://doi.org/10.1146/annurev-ecolsys-121415-032116>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.
- Felden, A., Paris, C. I., Chapple, D. G., Haywood, J., Suarez, A. V., Tsutsui, N. D., ... Gruber, M. A. M. (2018). Behavioural variation and plasticity along an invasive ant introduction pathway. *Journal of Animal Ecology*, 87, 1653–1666. <https://doi.org/10.1111/1365-2656.12886>
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, 27, 2215–2233. <https://doi.org/10.1111/mec.14584>
- Gautier, M. (2015). Genome-Wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, 201, 1555–1579. <https://doi.org/10.1534/genetics.115.181453>
- Goubert, C., Henri, H., Minard, G., Valiente Moro, C., Mavingui, P., Vieira, C., & Boulesteix, M. (2017). High-throughput sequencing of transposable element insertions suggests adaptive evolution of the invasive Asian tiger mosquito towards temperate environments. *Molecular Ecology*, 26, 3968–3981. <https://doi.org/10.1111/mec.14184>
- Goudet, J. (2017). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, 86, 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Han, J., Nalam, V. J., Yu, I. C., & Nachappa, P. (2019). Vector competence of thrips species to transmit soybean vein necrosis virus. *Frontiers in Microbiology*, 10, 431. <https://doi.org/10.3389/fmicb.2019.00431>
- Hedrick, P. W., & Parker, J. D. (1997). Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annual Review of Ecology and Systematics*, 28, 55–83. <https://doi.org/10.1146/annurev.ecolsys.28.1.55>
- Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *Bmc Evolutionary Biology*, 8, 289–289. <https://doi.org/10.1186/1471-2148-8-289>
- Hill, M. P. (2013). A predicted niche shift corresponds with increased thermal resistance in an invasive mite, *Halotydeus destructor*. *Global Ecology and Biogeography*, 22, 942–951.
- Hodgson, C. J., Abbas, G., Arif, M. J., Saeed, S., & Karar, H. (2008). *Phenacoccus solenopsis* Tinsley (Sternorrhyncha: Coccoidea: Pseudococcidae), an invasive mealybug damaging cotton in Pakistan and India, with a discussion on seasonal morphological variation. *Zootaxa*, 1913, 1–35. <https://doi.org/10.11646/zootaxa.1913.1.1>
- Hoffmann, A. A. (2017). Rapid adaptation of invertebrate pests to climatic stress? *Current Opinion in Insect Science*, 21, 7–13. <https://doi.org/10.1016/j.cois.2017.04.009>
- Ibrahim, S. S., Moharum, F. A., & Elghany, N. M. A. (2015). The cotton mealybug *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae) as a new insect pest on tomato plants in Egypt. *Journal of Plant Protection Research*, 55, 48–51. <https://doi.org/10.1515/jppr-2015-0007>
- Jaime, H. C., Kristoffer, F., Pedro, C. L., Damian, S., Juhl, J. L., Christian, V. M., & Peer, B. (2017). Fast genome-wide functional annotation

- through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution*, 34, 2115–2122.
- Jarošík, V., Kenis, M., Honěk, A., Skuhrovec, J., & Pyšek, P. (2015). Invasive insects differ from non-invasive in their thermal requirements. *PLoS ONE*, 10, e0131072. <https://doi.org/10.1371/journal.pone.0131072>
- Jensen, J. D., Foll, M., & Bernatchez, L. (2016). The past, present and future of genomic scans for selection. *Molecular Ecology*, 25, 1–4. <https://doi.org/10.1111/mec.13493>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Knaus, B. J., & Grünwald, N. J. (2017). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17, 44–53.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357. <https://doi.org/10.1038/nmeth.1923>
- Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, 17, 386–391. [https://doi.org/10.1016/S0169-5347\(02\)02554-5](https://doi.org/10.1016/S0169-5347(02)02554-5)
- Liebold, A., & Tobin, P. (2008). Population ecology of insect invasions and their management. *Annual Review of Entomology*, 53, 387–408. <https://doi.org/10.1146/annurev.ento.52.110405.091401>
- Lombaert, E., Guillemaud, T., Lundgren, J., Koch, R., Facon, B., Grez, A., ... Rhule, E. (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, 23, 5979–5997.
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23, 2178–2192.
- Lu, Y. Y., Zeng, L., Wang, L., Xu, Y. J., & Chen, K. W. (2008). Precaution of solenopsis mealybug *Phenacoccus solenopsis* Tinsley. *Journal of Environmental Entomology*, 30, 386–387.
- Ma, L., Cao, L. J., Gong, Y. J., Hoffmann, A. A., Zeng, A. P., Wei, S. J., & Zhou, Z. S. (2019). Development of novel microsatellites for population genetic analysis of *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae) based on genomic analysis. *International Journal of Biological Macromolecules*, 121, 1135–1144. <https://doi.org/10.1016/j.ijbiomac.2018.10.143>
- Munshi-South, J., Zolnik, C. P., & Harris, S. E. (2016). Population genomics of the Anthropocene: Urbanization is negatively associated with genome-wide variation in white-footed mouse populations. *Evolutionary Applications*, 9, 546–564. <https://doi.org/10.1111/eva.12357>
- Normark, B. B. (2004). Haplodiploidy as an outcome of coevolution between male-killing cytoplasmic elements and their hosts. *Evolution*, 58, 790–798. <https://doi.org/10.1111/j.0014-3820.2004.tb00412.x>
- Park, S. (2001). *The excel microsatellite toolkit (version 3.1)*. Dublin, Ireland: Animal Genomics Laboratory, UCD.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7, e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Piry, S., Luikart, G., & Cornuet, J. (1999). Computer note. BOTTLENECK: A computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity*, 90, 502–503. <https://doi.org/10.1093/jhered/90.4.502>
- Prasad, Y. G., Prabhakar, M., Sreedevi, G., Rao, G. R., & Venkateswarlu, B. (2012). Effect of temperature on development, survival and reproduction of the mealybug, *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae) on cotton. *Crop Protection*, 39, 81–88. <https://doi.org/10.1016/j.cropro.2012.03.027>
- Pudovkin, A., Zaykin, D., & Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, 144, 383–387.
- Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jonsson, B., ... Hansen, M. M. (2014). Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, 23, 2514–2528. <https://doi.org/10.1111/mec.12753>
- Pujolar, J. M., Jacobsen, M. W., Frydenberg, J., Als, T. D., Larsen, P. F., Maes, G. E., ... Hansen, M. M. (2013). A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. *Molecular Ecology Resources*, 13, 706–714. <https://doi.org/10.1111/1755-0998.12117>
- Rašić, G., Filipović, I., Weeks, A. R., & Hoffmann, A. A. (2014). Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector. *Aedes Aegypti*. *BMC Genomics*, 15, 275. <https://doi.org/10.1186/1471-2164-15-275>
- Raymond, M., & Rousset, F. (1995). GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249. <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
- Reilstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24, 4348–4370. <https://doi.org/10.1111/mec.13322>
- Renault, D., Laparie, M., McCauley, S. J., & Bonte, D. (2018). Environmental adaptations, ecological filtering, and dispersal central to insect invasions. *Annual Review of Entomology*, 63, 345–368. <https://doi.org/10.1146/annurev-ento-020117-043315>
- Roura-Pascual, N., Hui, C., Ikeda, T., Leday, G., Richardson, D. M., Carpintero, S., ... Worner, S. P. (2011). Relative roles of climatic suitability and anthropogenic influence in determining the pattern of spread in a global invader. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 220–225. <https://doi.org/10.1073/pnas.1011723108>
- Seabra, S. G., Bras, P. G., Zina, V., Borges da Silva, E., Rebelo, M. T., Figueiredo, E., ... Franco, J. C. (2013). Molecular evidence of polyandry in the citrus mealybug, *Planococcus citri* (Hemiptera: Pseudococcidae). *PLoS ONE*, 8, e68241. <https://doi.org/10.1371/journal.pone.0068241>
- Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., ... Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435. <https://doi.org/10.1038/ncomms14435>
- Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, 68, 1–15.
- Stoeckel, S., Grange, J., Fernandez-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Molecular Ecology*, 15, 2109–2118.
- Tinsley, J. (1898). An ants'-nest coccid from New Mexico. *The Canadian Entomologist*, 30, 47–48. <https://doi.org/10.4039/Ent3047-2>
- Trucchi, E., Gratton, P., Whittington, J. D., Cristofari, R., Le, M. Y., Stenseth, N. C., & Le, B. C. (2014). King penguin demography since the last glaciation inferred from genome-wide data. *Proceedings of the Royal Society B-Biological Sciences*, 281, 296–306. <https://doi.org/10.1098/rspb.2014.0528>
- Urbanski, J., Mogi, M., O'Donnell, D., DeCotiis, M., Toma, T., & Armbruster, P. (2012). Rapid adaptive evolution of photoperiodic response during invasion and range expansion across a climatic gradient. *The American Naturalist*, 179, 490–500. <https://doi.org/10.1086/664709>
- van Boheemen, L. A., Atwater, D. Z., & Hodgins, K. A. (2019). Rapid and repeated local adaptation to climate in an invasive plant. *New Phytologist*, 222, 614–627. <https://doi.org/10.1111/nph.15564>
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources*, 4, 535–538.

- Vanwallendael, A., Hamann, E., & Franks, S. J. (2018). Evidence for plasticity, but not local adaptation, in invasive Japanese knotweed (*Reynoutria japonica*) in North America. *Evolutionary Ecology*, 32, 1–16. <https://doi.org/10.1007/s10682-018-9942-7>
- Wang, Y. P., Gillianw, W., & Zhang, R. Z. (2010). The potential distribution of an invasive mealybug *Phenacoccus solenopsis* and its threat to cotton in Asia. *Agricultural and Forest Entomology*, 12, 403–416. <https://doi.org/10.1111/j.1461-9563.2010.00490.x>
- Wei, J. F., Zhang, H. F., Zhao, W. Q., & Zhao, Q. (2017). Niche shifts and the potential distribution of *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae) under climate change. *PLoS ONE*, 12, e0180913. <https://doi.org/10.1371/journal.pone.0180913>
- Wei, S. J., Cao, L. J., Gong, Y. J., Shi, B. C., Wang, S., Zhang, F., ... Chen, X. X. (2015). Population genetic structure and approximate Bayesian computation analyses reveal the southern origin and northward dispersal of the oriental fruit moth *Grapholita molesta* (Lepidoptera: Tortricidae) in its native range. *Molecular Ecology*, 24, 4094–4111.
- Willoughby, J. R., Harder, A. M., Tennessen, J. A., Scribner, K. T., & Christie, M. R. (2018). Rapid genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. *Molecular Ecology*, 4041–4051. <https://doi.org/10.1111/mec.14726>
- Wu, S. A., & Zhang, R. Z. (2009). A new invasive pest, *Phenacoccus solenopsis*, threatening seriously to cotton production. *Chinese Bulletin of Entomology*, 1, 159–162.
- Ying, J. I., Rong, J. I., & Huang, R. X. (2004). Invasive Species -*Agrilus mali* Matsumura and damage in Xinjiang. *Xinjiang Agricultural Sciences*, 41, 31–33.
- Zhu, Y. Y., Huang, F., & Lu, Y. B. (2011). Bionomics of mealybug *Phenacoccus solenopsis* Tinsley (Hemiptera:Pseudococcidae) on cotton. *Acta Entomologica Sinica*, 54, 246–252.

## BIOSKETCH

This work is a part of the Master study of Ling Ma on invasion genetics of the cotton mealybug *Phenacoccus solenopsis*. The research team includes specialists working on biology, ecology, taxonomy, population genetics and management of the *P. solenopsis*.

Author contributions: Z.S.Z. and S.J.W. conceived and designed the study; L.M. and L.J.C. conducted the molecular works; L.M., S.J.W. and L.J.C. analysed the data; Z.S.Z., S.J.W., Y.J.G., J.C.C., A.P.Z., X.B.W. and H.H.C. organized the collection of specimens; L.M., S.J.W., A.A.H. and Z.S.Z. discussed the results. L.M., S.J.W. and A.A.H. wrote the paper.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ma L, Cao L-J, Hoffmann AA, et al.

Rapid and strong population genetic differentiation and genomic signatures of climatic adaptation in an invasive mealybug. *Divers Distrib*. 2020;26:610–622. <https://doi.org/10.1111/ddi.13053>