Microsatellite-based analysis of genetic structure and gene flow of *Mythimna separata* (Walker) (Lepidoptera: Noctuidae) in China

Mei-Mei Li | Bo-Liao Li | Shi-Xiong Jiang | Yu-Wan Zhao | Xiang-Li Xu | Jun-Xiang Wu

State Key Laboratory of Crop Stress Biology in Arid Areas, College of Plant Protection, Northwest A & F University, Yangling, China

**Correspondence**
Xiang-Li Xu and Jun-Xiang Wu, State Key Laboratory of Crop Stress Biology in Arid Areas, College of Plant Protection, Northwest A & F University, No. 3 Taicheng Road, 712100 Yangling, Shaanxi, China. Emails: xuxiangli@nwsuaf.edu.cn (X-L X) and junxw@nwsuaf.edu.cn (J-X W)

**Funding information**
the National Key R & D Program of China, Grant/Award Number: 2017YFD0201807; Agro-scientific Research in the Public Interest of China, Grant/Award Number: 201403031

**Abstract**
The oriental armyworm, *Mythimna separata*, is a serious agricultural pest in China. Seasonal and roundtrip migration has recently led to sudden, localized outbreaks and crop losses. To evaluate genetic differentiation between populations in eastern and western China and elucidate gene flow, the genetic structure of 20 natural populations from nine provinces was examined using seven microsatellite markers. The results indicated high genetic diversity. However, little to moderate ($0 < F_{ST} < 0.15$) genetic differentiation was detected, and there was no correlation between genetic distance and geographical distance. Bayesian clustering analysis identified three groups whereas discriminant analysis of principal components identified ten clusters that were considered as two clear-cut clusters and one admixed group. Gene flow occurred frequently in most population pairs, and an asymmetrical migration rate was detected in several pairwise population comparisons. The bottleneck test showed that few populations had experienced recent bottlenecks. Correspondingly, large-scale and long-distance migration of *M. separata* has caused low genetic differentiation and frequent gene exchange. Our findings are important for studying genetic evolution and help to improve predictions of *M. separata* outbreaks in China.

**KEYWORDS**
gene flow, genetic diversity, microsatellites, *Mythimna separata*, population structure

1 | INTRODUCTION

*Mythimna separata* (Walker) (Lepidoptera: Noctuidae) is a common migratory pest in Eastern Asia and some parts of Oceania (Chen, Sun, Wang, Zhai, & Bao, 1995; Farrow & McDonald, 1987). Outbreaks occur intermittently in most regions of China (Zeng, Jiang, & Liu, 2013). It preferentially feeds on Graminaceous crops, such as maize (*Zea mays* L.), wheat (*Triticum aestivum* L.), and rice (*Oryza sativa* L.). The gluttonous behavior of larvae, their short life cycle (Figure 1), and multiple generations per year lead *M. separata* to inflict serious damages and economic losses (Jiang, Zhang, Chen, & Luo, 2014a). Thus, it presents a serious threat to crop production throughout the growing season (Jiang, Jiang, Zhang, Cheng, & Luo, 2014b). The second and third generations of *M. separata* present the most serious infestations, whereas with the third generation becoming the predominant one to produce local infestations in China in recent years...
The armyworm moth undertakes long-range and multigeneration roundtrip migration between southern and northern China from March to mid-September each year (Jiang, Luo, Zhang, Sappington, & Hu, 2011; Li, Wang, & Hu, 1964; Zhang, Zhang, Li, Jiang, & Zeng, 2013). This makes it difficult to predict population dynamics and formulate proper management decisions. Its overwintering boundary indicates that *M. separata* is unable to overwinter in northern China and no diapause period occurs during development (Li, 1993). Therefore, large-scale outbreaks of *M. separata* in northwestern China are probably attributed to immigration from the eastern regions (Cheng & Zhao, 2016; Zhou et al., 2017). However, seasonal dynamics of armyworm movement between eastern and western China remain unclear.

Based on isozyme analyses, population genetic studies of *M. separata* concluded that oriental armyworm populations formed a large panmictic population from five localities in China (Hao, Li, & Lin, 1992). Researchers further confirmed high genetic variation among individuals of studied populations, and little genetic differentiation was found among geographically distinct populations in China. This is likely due to increased gene flow as a result of their long-distance migration (Jiang, Luo, & Zhang, 2007; Li, Li, Wu, & Xu, 2018). Furthermore, intersimple sequence repeat (ISSR) analysis implied that highly frequent gene flow among 23 geographically separated armyworm populations from various latitudes prevented genetic differentiation normally caused by genetic drift (Chen et al., 2017). However, their conclusions were unable to reveal population genetic structure and evolutionary history due to limited sampling and the use of molecular markers of low reproducibility or diversity (Chen et al., 2017; Hao et al., 1992; Jiang et al., 2007).

Many studies have been conducted on the migratory flight, northern boundary of overwintering, and evolutionary history of armyworms (Chen et al., 1989, 2017; Hao et al., 1992; Jiang et al., 2007; Li, 1993; Li et al., 1964; Yin, Feng, Cheng, & Cao, 2003; Zhang et al., 2013). However, a comprehensive study of population genetic variation and population structure in China has not been accomplished. The present study analyzed the population structure of 20
geographically distinct populations of the oriental armyworm collected from nine provinces in China, using microsatellite markers. Knowledge about genetic structure among subpopulations, population differentiation, and gene flow of *M. separata* in China is greatly important for determining migration patterns, and thus, the source of outbreaks; it also provides a theoretical framework for prediction and determines control strategies to prevent outbreaks.

### 2 | MATERIALS AND METHODS

#### 2.1 | Sampling collection and DNA extraction

The oriental armyworm larvae samples were collected from maize (*Z. mays*) fields from nine provinces at 20 geographical sites in China (Figure 1; Table 1). A total of 542 armyworms were sampled during the maize-growing season from May 2016 to August 2017. Each armyworm was collected at a distance of at least 50 m from any other sampled individuals, to avoid sampling siblings. Armyworm larvae were preserved in absolute ethanol and stored at −20°C. Total genomic DNA for genetic analysis was extracted from whole larvae using Biospin insect genomic DNA extraction Kit (Bioer Technology Co., Ltd).

#### 2.2 | PCR amplification and microsatellite genotyping

Seven polymorphic microsatellite loci, A1-91, A3-81, JGTT29, MS1, MS2, MS3, and MS4, developed for *M. separata* were scored for genotyping (Li et al., 2018; Zhang & Zhai, 2008). The three primers used were as follows: a sequence-specific forward primer linked with M13-tail at its 5′ end, a sequence-specific reverse primer, and a FAM or HEX fluorescent-labeled M13 (−21) primer (Schuelke, 2000). All primers were synthesized by AuGCT Biotechnology. PCR reactions were performed with a total volume of 20 µl, containing 10 µl Golden Star T6 Super PCR Mix (1.1×), 0.4 µl forward primer (10 µM), 1.6 µl fluorescent-labeled primer (FAM or HEX) and reverse primer (10 µM), respectively, 1.2 µl genomic DNA (20–50 ng/µl), and 5.2 µl ddH₂O. Amplifications were initiated with a predegeneration 94°C for 2 min, followed by 30 amplification cycles consisting of 94°C for 10 s, 15 s at the primer-specific annealing temperature, 72°C for 15 s. This was followed by 10 cycles including 94°C for 15 s, 53°C for 20 s, and 72°C for 20 s, and ending with 72°C for 5 min for a final extension. The primer-specific annealing temperature of each primer was described in previous report (Li et al., 2018). The PCR products of the target band were genotyped by capillary electrophoresis with two-color fluorescent

### TABLE 1  Sampling information of *Mythimna separata* larvae in China

| Collection region (province) | Population code (PC) | Geographical population | Sample size | Generation a | Collection time | Latitude/Longitude |
|-----------------------------|----------------------|-------------------------|-------------|--------------|----------------|-------------------|
| Nonoverwintering regions    |                      |                         |             |              |                |                   |
| Henan (HN)                  | HNX                  | Gushi, Xinyang          | 22          | 1            | 3-May-16       | 32.193°N/115.589°E |
|                            | HNL                  | Yichuan, Luoyang        | 30          | 3            | 3-Aug-16       | 34.401°N/112.642°E |
|                            | HNH                  | Linying, Luohu          | 29          | 2            | 22-Jun-17      | 33.739°N/113.966°E |
| Shanxi (SY)                 | SYC                  | Ruicheng, Yuncheng      | 28          | 3            | 8-Aug-16       | 34.694°N/110.840°E |
| Shaanxi (SX)                | SXH                  | Huayin, Weinan          | 29          | 3            | 5-Aug-16       | 34.546°N/110.075°E |
|                            | SXF                  | Fuping, Weinan          | 28          | 2            | 1-Jul-17       | 34.970°N/109.242°E |
|                            | SXM                  | Mian county, Hanzhong   | 27          | 2            | 2-Jul-17       | 33.075°N/106.707°E |
|                            | SXX                  | Xingping, Xianyang      | 27          | 2            | 3-Jul-17       | 34.303°N/108.468°E |
|                            | SXB                  | Qianyang, Baoji         | 25          | 2            | 4-Jul-17       | 34.832°N/107.284°E |
|                            | SXL                  | Luochuan, Yanan         | 28          | 3            | 4-Aug-17       | 35.771°N/109.429°E |
|                            | SXW                  | Heyang, Weinan          | 26          | 3            | 7-Aug-17       | 35.216°N/110.236°E |
| Jilin (JL)                  | JLC                  | Nanguan, Changchun      | 26          | 2            | 23-Jun-17      | 43.811°N/125.412°E |
|                            | JLS                  | Gongzhuling, Siping     | 28          | 2            | 24-Jun-17      | 43.505°N/124.823°E |
| Hebei (HB)                  | HBB                  | Xinshi, Baoding        | 22          | 3            | 8-Aug-17       | 38.956°N/115.450°E |
|                            | HBC                  | Yunhe, Changzhou        | 26          | 3            | 10-Aug-17      | 38.283°N/116.820°E |
| Liaoning (LN)               | LNF                  | Zhangwu, Fuxin          | 25          | 3            | 9-Aug-17       | 42.735°N/122.567°E |
| Shandong (SD)               | SDJ                  | Changqing, Jinan        | 29          | 3            | 10-Aug-17      | 36.458°N/116.704°E |
|                            | SDW                  | Wangtuan, Weihai        | 29          | 3            | 13-Aug-17      | 37.368°N/121.991°E |
| Inner mongolia (IM)         | NME                  | Hangjinqi, Ordos        | 28          | 3            | 11-Aug-17      | 40.271°N/107.041°E |
| Overwintering region        |                      |                         |             |              |                |                   |
| Yunnan (YN)                 | YNZ                  | Zhenxiong, Zhaotong     | 30          | 2            | 12-Jul-17      | 27.408°N/104.980°E |

*a* Presents the local occurrence generation.
2.3 | Genotypic analysis

2.3.1 | Levels of genetic diversity

Null allele frequency of each microsatellite locus was estimated using FreeNA (Chapuis & Estoup, 2007) following the Expectation-maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977), with a significance estimated by 1,000 bootstrap replicates. The ENA (excluding null alleles) correction method (Chapuis & Estoup, 2007) was used to estimate Weir’s (1996) global $F_{ST}$. The polymorphism information content (PIC) was calculated using PowerMarker 3.25 (Lu & Muse, 2005). The mean number of alleles (Na), effective alleles number (Ne), Shannon’s index (I), observed heterozygosity (Ho), and expected heterozygosity (He) were estimated using GenAlEx 6.5 (Peakall & Smouse, 2012) as measures of genetic variability within populations. Gene diversity (H), allelic richness (Ar), and inbreeding coefficient (Fis) within populations were evaluated using Fstat 2.9.3 (Goudet, 1995; Nei, 1973). Deviation from Hardy–Weinberg equilibrium (HWE) for each locus was tested, using Fisher’s exact tests implemented in GENEPOP 4.6 (Rousset, 2008). Linkage disequilibrium (LD) between pairs of loci was performed in Arlequin 3.5 with 10,000 permutations (Excoffier & Lischer, 2010). 

2.3.2 | Genetic structure and population differentiation

The population genetic variance was further analyzed by a global analysis of molecular variance (AMOVA) performed using Arlequin 3.5, and average F-statistics, pairwise differentiation ($F_{ST}$), and its significance levels (p-values) were estimated with 1,000 permutations (Excoffier & Lischer, 2010; Weir & Cockerham, 1984). We tested for isolation by distance by regressing pairwise genetic distance ($F_{ST}/(1 - F_{ST})$) against the natural logarithm of geographical distance (km) across populations for microsatellite loci using the Mantel test (Mantel, 1967) with 1,000 permutations of ZT software package (Bonnet & Van der Peer, 2002). Population structure was analyzed based on Bayesian model-based clustering using STRUCTURE 2.3.4 with admixture ancestry and correlated allele frequency model (Pritchard, Stephens, & Falush, 2007). Twenty independent runs for each testing genetic clusters (K) values which ranged from 1 to 20 were performed with a burn-in period of 50,000 iterations and 1,000,000 Markov Chain Monte Carlo (MCMC) replicates. The most likely number of clusters (K) was determined by considering log-likelihood values of each K and the Delta K method described by Evanno, Regnaut, and Goudet (2005), implemented in Structure Harvester (Earl & vonholdt, 2012). CLUMPP 1.1 was used for model averaging of individual ancestry coefficients across the 20 independent runs (Jakobsson & Rosenberg, 2007). Then clusters were visualized using Distruct 1.1 (Rosenberg, 2004). Further, discriminant analysis of principal components (DAPC) was used to identify subpopulations of the species using the Adegenet package for the R 3.5.3 (Jombart, 2008; Jombart, Devillard, & Balloux, 2010; R Core Team, 2019). The number of clusters was assessed using the function find.clusters, evaluating a range from 1 to 20. The optimal number of clusters was identified based on the Bayesian information criterion (BIC), as suggested by Jombart et al. (2010). The function find.clusters was used to transform original data into principal components (PC), retaining 200 PCs in the analysis. The dapc function performed a discriminant analysis using 50 PCs (>80% of variance explained), and nine eigenvalues were retained and examined. The assign.per.pop function was used to evaluate the proportions of successful reassignment of individuals to their original clusters. Specifically, we applied the a.score and a.optim.score functions were used to identify the optimal number of principal components to be retained. Twenty independent runs were executed and the average results were plotted.

2.3.3 | Analysis of gene flow and migration rate

Wright’s method was employed to estimate gene flow ($N_{m}$) from genetic data based on the relationship $F_{ST} = 1/(4N_{e}m + 1)$ (where $N_{e}$ is the effective size of each population and $m$ is the migration rate between populations; Slatkin, 1987; Wright, 1931). Bayesian inference for estimating rates of recent migration (over the last several generations) among populations based on multilocus genotypes using BayesAss 3.0 that relies on Markov chain Monte Carlo methods (Wilson & Rannala, 2003).

2.3.4 | Bottleneck analysis

The infinite allele model (IAM), the strict stepwise mutation model (SMM) and two-phase model (TPM) were applied with 10,000 simulation iterations to identify whether populations have experienced bottleneck effects in history under the assumption of mutation–drift equilibrium using BOTTLENECK 1.2.02 (Piry, Luikart, & Cornuet, 2001). A qualitative descriptor of allele frequency distribution (“mode-shift” indicator), which discriminates between bottlenecked and stable populations, was also utilized (Luikart, Allendorf, Sherwin, Cornuet, & Sherwin, 1998). Under the three models, the standardized differences test was removed from this study because the test is only used when approximately 20 or more polymorphic loci are available. The Wilcoxon signed-rank test, a more appropriate and powerful method, was applied to detect whether there was a significant heterozygote excess in each population.

3 | RESULTS

3.1 | Genetic diversity, Hardy–Weinberg, and linkage testing

A total of 193 alleles were detected over seven microsatellite loci, with a low null allele frequency ranging from 0.029 to 0.185.
Therefore, no loci were removed from further genetic analyses. Global $F_{ST}$ estimates, with and without ENA correction, were 0.0447 and 0.0453, respectively. The polymorphism information content (PIC) ranged from 0.628 to 0.935 (Table 2). For the population level, the mean number of alleles ($Na$) and effective number of alleles ($Ne$) were 9.5 and 5.23, respectively. The estimation for the average of Shannon’s index ($I$) was 1.808. The allelic richness ($Ar$) ranged from 6.499 to 11.449 per population, and gene diversity ranged from 0.712 to 0.828. Observed heterozygosity ($Ho$) ranged from 0.580 to 0.702, while expected heterozygosity ($He$) ranged from 0.694 to 0.810. The inbreeding coefficient ($Fis$) within individuals, relative to the rest of their subpopulation, was between 0.097 and 0.297 (Table 3). Additionally, the HBB and YNZ populations possessed the lowest and highest genetic diversity, respectively. Deviations from Hardy–Weinberg equilibrium (HWE) were apparent in six of 20 populations, revealing that these populations tended toward
heterozygote excess. Exact tests for departure from HWE across all loci suggested that 95 of the 140 locus-population combinations deviated significantly \((p < .05)\). Linkage disequilibrium presented in 63 of the 147 pairs of microsatellite loci, and no complete linkage of a certain locus was detected.

3.2 Population genetic structure

According to the Bayesian analysis performed with STRUCTURE 2.3.4, individuals were grouped into three main clusters (Figure 2a; Figure S1A,B). Although there was a clear signal of population structure, the optimal value of \(K\) was difficult to determine as no single unifying characteristic was apparent for any of the inferred groups at \(K = 3\). Generally, among them, cluster 1 dominated in the LNF and SXB populations, cluster 2 dominated in the JLC and YNZ populations, and cluster 3 dominated in the HBB and SYC populations (cluster with an estimated membership >0.800). However, the other 14 populations showed some level of admixture, and half of them did not completely form their own cluster (<0.500; Tables S1 and S2). Four populations (HNX, HNL, SXH, and SYC) sampled in 2016 shared a similar genetic structure (Figure 2b), which implied that their parent populations came from the same overwintering areas. Populations in the same generation had a similar genetic structure (Figure 2c), indicating the same parental
insect sources. The genetic structure of the oriental armyworm populations did not vary significantly per geographic divisions, and the YNZ population from overwintering region had a highly similar genetic structure to that of the JLC population from a nonoverwintering region (Figure 2d).

The Bayesian information criteria (BIC) curve in the DAPC analysis supported 10 clusters (Figure S2). Clusters 2, 4 and 10 were clearly differentiated, which indicated that LNF, SDW, and HNL belonged to disparate subgroups. Individuals in the other seven clusters had no distinct division (Figure 3). NME and SXB had similar genetic structure, and SDJ and HBB also had similar genetic structure but were differed from the structure of NME and SXB (Table S3). Inspection of the DAPC plot also revealed the presence of a genetic structure within these populations and was more likely to divide them into three groups (Figure 3). Per Bayesian clustering, no clear associations between the groups were identified using DAPC.

Combining results from Bayesian analysis in STRUCTURE and DAPC, individuals from NME and SXB to be genetically differentiated from LNF, as well as HBC (Figures 2 and 3). Although the four populations from 2016 were in various subgroups, they were weakly differentiated. Individuals from the JLC and YNZ populations belonged to different clusters and were partial overlapped in DAPC analysis despite they divided into multiple genetic clusters, while they had a similar genetic partitioning in Bayesian analysis. Genetic clustering of individuals in each population presented mixed ancestry (Figure 2 and 3), which led to multiple genetic clusters found in armyworm populations. Consequently, these armyworm populations were considered to originate from three differentiated subpopulations.

3.3 | Population genetic differentiation

The global analysis of molecular variance (AMOVA) test results attributed 4.54%, 19.72%, and 75.74% of the variation to differences among populations, among individuals within populations, and within individuals, respectively, based on the pairwise differences distance method (Table 4). The global $F_{ST}$ value, calculated across all populations and loci, was 0.045 ($p < .001$). It is generally assumed that $F_{ST} < 0.05$, $0.05 < F_{ST} <0.15$, $0.15 < F_{ST} <0.25$, and $F_{ST} > 0.25$ exhibit little, moderate, significant, and high levels of genetic differentiation, respectively (Wright, 1978). The pairwise $F_{ST}$ values (Table S4) were low to moderate, with a maximum between SXB and HBB (0.1129), which indicated a high level of movement among populations. Overall, 58.42% of populations demonstrated little genetic differentiation ($F_{ST} < 0.05$). With the exception of SDJ population, moderate differentiation presented between the HBB population and 18 the other populations ($F_{ST} > 0.05$). The findings illustrated that genetic divergences in the armyworm were mainly between individuals and that gene flow among populations in China was not limited.

The Mantel test revealed a positive but not significant correlation ($r = .154$, $p = .118$) between the $F_{ST}/(1 − F_{ST})$ ratio for pairs of populations against the natural logarithm-transformed geographical distance (Tables S4 and S5), demonstrating that the population genetic structure of $M. separata$ in China did not conform to the isolation by distance model (Figure 4). That is, no correlation between pairwise genetic distance and geographical distance of armyworm populations was evidenced. Furthermore, geographical distance did not reflect a barrier to gene flow.

3.4 | Gene flow and migration rate

Observations of the effective number of migrants per generation $(N_e m)$ of several population pairs were greater than 4 (Table S5), indicating that active dispersal ability among those populations was not limited. It is generally assumed that $N_e m$ equal to 1 is sufficient to bring the two populations genetically closer to each other, and if $N_e m > 4$, then local populations belong to one panmictic population (Wright, 1938). The value of $N_e m$ between SXH and SYC populations...
was 2.69239, indicating extremely frequent gene exchange or perhaps the same parent population. The gene flow analysis revealed a long-distance dispersal event (especially in SXX/YNZ and JLS/SDJ) during a period of large-scale migration from June to August 2017. Theoretically, \( N_e \) also assumes that populations are in migration-drift equilibrium, which occurs when the increase in genetic differentiation caused by drift is equal to the decrease in genetic differentiation caused by gene flow (Joanna, Heather, & Stephen, 2011). Estimates of gene flow based on microsatellite data analysis indicated high levels of gene flow and the absence of genetic differentiation among population pairs.

The results showed that the recent migration rates ranged from 0.0059 to 0.1817, and asymmetric gene flow existed in some interpopulations, especially JLS, SXH, and SXX (Table S6). The rate of self-distribution within the SXH population was the highest (0.8613), while that of the SDJ population was the lowest (0.6733). In 2016, the analysis for migration direction across four populations (HNL, HNX, SXH, and SYC) confirmed that the Henan populations (HN) migrated westward to Shaanxi and Shanxi regions (SX and SY) of China. Moreover, Yunnan populations (YN) migrated eastward or northward to Shandong regions (SD) and near Jilin (JL); the Henan population (HN) migrated westward to Shaanxi (SX) and farther northward to Inner Mongolia (IM) in 2017 (Table 1; Figure 1).

### 3.5 Population demography

An attempt to test heterozygote excess was inconsistent under the IAM, TPM, and SMM models (Table 5). The bottleneck test indicated that heterozygote excess was not significant (\( p > .05 \)) in any populations under the SMM model. Conversely, heterozygote excess of most populations was apparent (\( p < .05 \)) under the IAM model, except HBC and JLS populations, elucidating that those populations experienced a reduction in effective size. The discrepancy between IAM and SMM might result from either a small sample size or an influence of compound and complex loci. Under TPM model, significant heterozygote excess (\( p < .05 \)) was detected in three of the populations (HBB, HNX, and SXH), indicating that these populations deviated from mutation–drift equilibrium and might have recently undergone a genetic bottleneck; allele frequency presented a normal L-shaped distribution (Table 5).

### 4 DISCUSSION

#### 4.1 Genetic diversity

In this study, 193 alleles were found and *M. separata* had a higher genetic diversity based on seven polymorphic microsatellites with low null allele frequency (Tables 2 and 3). The \( F_{st} \) values for all populations were greater than zero, indicating a heterozygote deficiency, which caused a deviation from HWE (Table 3). Overall, \( H_o \) values for all populations were lower than \( H_e \), which might be attributed to null alleles, Wahlund effect, and the population not being in HWE (Joanna et al., 2011). There was no complete linkage of any locus, indicating that the selected loci were evenly distributed in the genome of armyworm and were relatively independent in the process of generational transmission (Torriani, Mazzi, Hein, & Dorn, 2010). The genetic indices showed that the level of genetic diversity in most populations was similar. Higher genetic diversity was observed in the second and third generations, except the HBB population. Furthermore, the YNZ population from the overwintering region exhibited a higher genetic diversity than others from nonoverwintering areas. Abundant genetic diversity was shown to be the genetic

### TABLE 4 Analysis of molecular variance on seven microsatellite loci in geographically distinct populations of *Mythimna separata*

| Source of variation | \( df \) | Sum of squares | Variance components | Percentage variation | \( F \)-statistics |
|--------------------|--------|----------------|---------------------|---------------------|------------------|
| Among populations  | 19     | 196.966        | 0.13180 Va          | 4.54                | \( F_{is} = 0.207 \) \( p < .001 \) |
| Among individuals within populations | 522 | 1,717.879     | 0.57237 Vb         | 19.72               | \( F_{ST} = 0.045 \) \( p < .001 \) |
| Within individuals | 542   | 1,176.500      | 2.19894 Vc          | 75.74               | \( F_{IT} = 0.243 \) \( p < .001 \) |
| Total              | 1,083 | 3,091.345      | 2.90311             |                     |                  |

**FIGURE 4** Scatter plots of pairwise genetic distance \( (F_{ST}/(1 - F_{ST})) \) against the natural logarithm of geographical distance (km) for pairwise population comparisons.
basis for its strong adaptive potential (Chen et al., 2017; Wang, Yang, Lu, Zhou, & Wu, 2017), which aids in understanding why *M. separata* localized, large-scale outbreaks occurred in China in recent decades (Cheng & Zhao, 2016; Jiang et al., 2011; Jiang, Zhang, et al., 2014a; Zeng et al., 2013).

### 4.2 Genetic structure

Bayesian clustering and DAPC analyses (Figures 2 and 3) suggested that the populations were mostly clustered into three groups. All populations were admixtures of these clusters, which could be only explained by the highly migratory ability of the armyworm gave rise to extensive gene flow. Although genetic structure appeared irrelevant to both geographical divisions and occurrence generations of *M. separata* populations throughout China, four populations from 2016 had a similar genetic partitioning which grouped in cluster 3 (Figure 2). As in Coulon et al. (2008), individuals with maximum inferred ancestry <0.6 were not assigned to any group and were considered to exhibit admixed assignation. That was, nine of the populations could not be assigned to any cluster (Table S1). Furthermore, DAPC confirmed that there was little genetic distinction among clusters 1, 3, and 5–9, which mainly contained 15 populations excluding: HNL, LNF, SDW, SXF, and SXW (Figure 3). Based on these results, we inferred that individuals from 20 geographic populations in the same cluster had the same immigration source; they likely shared a common gene pool in the past. AMOVA analysis found low to moderate genetic differentiation (global $F_{ST} = 0.045, p < .001$) based on $F_{ST}$ among populations. Our results indicated that migratory dispersal led to differences in population structure but they still belonged to a panmictic population owing to their high capacity for long-distance migration.

### 4.3 Genetic differentiation and migration

There was no significant correlation between geographical distance and genetic distance among any of the 20 populations (Figure 4). Our estimates (Table S5) suggested extensive gene flow ($N_e m > 1.0$), between pairwise populations, even between most adjacent populations of *M. separata*. This might be owing to its migratory behavior and nongeographical isolation. Therefore, a high gene flow of armyworms (Table S5) had a homogenizing effect on the genetic variation over geographic populations, counteracting random drift, selection, and mutation (Sun, Li, Yang, & Hong, 2012). Notably, the analyses of 20 geographically separated *M. separata* populations showed little to moderate genetic differentiation and a high gene flow (Tables S4 and S5). Insect species might be divided into genetically diverse populations that cover large and heterogeneous geographical regions with limited gene flow (Chen, Tan, Liu, Wang, & Li, 2000). In contrast, geographically distinct armyworm populations did not exhibit genetic differentiation, and isolation by distance was absent due to extensive gene flow, which further validated our findings. Previous research elucidated that frequent

| PC-sample size | Heterozygosity excess p-values | Distribution of allelic frequencies |
|----------------|-------------------------------|-----------------------------------|
|                | IAM  | TPM  | SMM  |                     |
| HBB-22         | .00391 | .03906 | .96094 | Normal L-shaped     |
| HBC-26         | .05469 | .98828 | 1.00000 |                     |
| HNH-29         | .00391 | .05469 | .98047 |                     |
| HNL-30         | .00781 | .71094 | .98047 |                     |
| HNX-22         | .00391 | .03906 | .59375 |                     |
| JLC-26         | .02734 | .65625 | 1.00000 |                     |
| JLS-28         | .05469 | .65625 | .97266 |                     |
| LNF-25         | .01172 | .76563 | .99609 |                     |
| NME-28         | .01172 | .53125 | 1.00000 |                     |
| SDJ-29         | .00781 | .81250 | 1.00000 |                     |
| SDW-29         | .02734 | .53125 | .98047 |                     |
| SXB-25         | .01953 | .53125 | .98828 |                     |
| SXF-28         | .03906 | .71094 | .98047 |                     |
| SXH-29         | .00391 | .02734 | .71094 |                     |
| SXL-28         | .00391 | .28906 | .97266 |                     |
| SXM-27         | .00391 | .18750 | .81250 |                     |
| SXW-26         | .00781 | .46875 | .99219 |                     |
| SXX-27         | .01953 | .76563 | 1.00000 |                     |
| SYC-28         | .00781 | .34375 | .98828 |                     |
| YNZ-30         | .00391 | .81250 | 1.00000 |                     |

Note: Population codes are given in Table 1; bold indicates significant at $p < .05$. Table 5: Wilcoxon’s signed-rank test for heterozygote excess with probability values based on three models (infinite allele model [IAM], two-phase model [TPM], and strict stepwise mutation model [SMM]) and distribution of allelic frequencies.
genetic exchange occurred between widely separated populations of *Helicoverpa zea* and that these populations did not show any signs of differentiation based on location (Perera & Blanco, 2011). The lack of genetic differentiation more likely reflected genetic mixing between eastern and western North American monarchs, Danaus plexippus (Lyons et al., 2012). Therefore, migratory insects form a large panmictic population experience enhanced gene flow because of genetic exchange, migratory flight, and long-range dispersal. Additionally, the present study articulated that there was no population differentiation caused by genetic drift and no clear evidence for isolation by distance, which is consistent with other studies of this insect (Chen et al., 2017; Li et al., 2018).

The migration rates in BayesAss suggested that migration was directional (Table S6). Thus, the parent populations of the northwestern outbreak areas were from northeastern and/or southern overwintering regions. Specifically, in northwestern China, insect source of the second generation mostly came from Henan, while the third generation migrated from northern China and/or Hebei. The results of migration rates were mainly consistent with pairwise $F_{ST}$ and $N_{m}$ values, that was, little genetic differentiation corresponded to extensive gene flow and large migration rates (Tables S4–S6). Furthermore, the Wilcoxon signed-rank test of heterozygote excess (Table 5) elucidated that the studied populations were not subjected to recent bottlenecks; hence, genetic diversity across populations was not affected.

Generally, there are two large-scale and long-distance migrations, that is, seasonal and roundtrip flight between southern and northern China (Li et al., 1964). According to the results of genetic structure, gene exchange, and migration rate, the oriental armyworm exhibited a particular flight pathway between eastern and western China. Armyworm moths in western China were from overwintering regions that migrated northward to HN in mid-April. Its offspring adults that migrated northwestward to the central SX plain, SY, and even possibly continued migrating northward to IM or migrated to north China plain (HB) and northeastern region (JL, LN, and HLJ) from late May to early June. In late July, most adults in northeastern region migrated southward to HB/SD or so far as to Henan. From late August, the armyworms in the northwestern areas (SX, SY, and IM) migrated southward back to overwintering regions (FJ, GD, and GX; Table 1; Figure 1). This migratory pathway agrees with the seasonal north–south migration pattern, based on mark-release-recapture data and trajectory simulation analysis of northeast populations from Shanxi, southern HB, northern HN, and most of the SD regions (Jiang et al., 2011; Li et al., 1964). The next generation migrated through the central and western regions of HB and IM, and some flew directly from the SD Peninsula across the Bohai Sea to the north of LN and center of JL by radar observation (Chen et al., 1989; Zhang et al., 2013). It was further observed that autumn migration of *M. separata* that originated in northeastern China (i.e., LN, JL, HLJ, and part of IM regions) could immigrate into east–central China (SD) and subsequently to southern China (i.e., FJ, GD, and GX; Table 1; Figure 1) within a week for overwintering (Feng et al., 2008). Considering the lack of data on the overwintering generation, more research is needed to elucidate the migratory patterns in the western areas, in order to better predict and prevent outbreaks.

5 | CONCLUSIONS

Microsatellite markers elucidated higher genetic diversity of the second and the third generations of *M. separata*. Our analysis showed that little population structure, the lack of genetic differentiation and its extensive gene flow among populations in China, which in consistent with the fact that populations of this species undertakes a long-distance and large-scale migration. In addition, bottleneck effects and genetic drift have not yet affected the genetic differentiation of *M. separata*. Finally, sampling including all generations occurring in a year across its entire distribution in China is required to further confirm genetic structure of populations, which reflects the interaction of genetic drift, mutation, and migration, with influences from life history. Greater knowledge on the genetic variation in natural populations of *M. separata* will influence regional outbreak patterns in China by ensuring proper pest prediction and control in advance.

ACKNOWLEDGMENTS

We especially thank associate Prof. Hui Dong of the Plant Protection College, Shenyang Agricultural University for his help with the collection of *Mythimna separata* samples in Liaoning. Funding for this study was provided jointly by Special Fund for Agro-scientific Research in the Public Interest of China (201403031) and the National Key R & D Program of China (2017YFD0201807).

CONFLICT OF INTEREST

None declared.

AUTHORS CONTRIBUTION

X.L.X. and J.X.W. conceived the study; M.M.L., S.X.J. and Y.W.Z. participated in collecting the biological samples; and M.M.L. and B.L.L. analyzed the data and wrote the manuscript. All authors contributed to editing the manuscript.

ORCID

Mei-Mei Li https://orcid.org/0000-0003-2035-0406

DATA AVAILABILITY STATEMENT

Microsatellite data file has been accessioned at Dryad (datadryad.org) under the title: “Microsatellite-based analysis of genetic structure and gene flow of Mythimna separata (Walker) (Lepidoptera: Noctuidae) in China” (https://doi.org/10.5061/dryad.5tb2rbp0n).
REFERENCES

Bonnet, E., & Van der Peer, Y. (2002). ZT: A software tool for simple and partial Mantel tests. Journal of Statistical Software, 7(10), 1–12.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution, 24(3), 621–631. https://doi.org/10.1093/molbev/msi191

Chen, R. L., Bao, X. Z., Drake, V. A., Farrow, R. A., Wang, S. Y., Sun, Y. J., & Zhai, B. P. (1989). Radar observations of the spring migration into northeastern China of the oriental armyworm, Mythimna separata and other insects. Ecological Entomology, 14(2), 149–162.

Chen, R. L., Sun, Y. J., Wang, S. Y., Zhai, B. P., & Bao, X. Y. (1995). Migration of the oriental armyworm Mythimna separata in East Asia in relation to weather and climate. I. Northeastern China. In A. G. Drake & V. A. Gatehouse (Eds.), Insect migration: Tracking resource in space and time (pp. 93–104). Cambridge, UK: Cambridge University Press.

Chen, X. Y., Li, Y. H., Hao, R. S., Ke, R., Li, Y. C., Zhao, S. P., & Gui, F. R. (2017). Genetic diversity of Mythimna separata (Walker) (Lepidoptera: Noctuidae) in Yunnan Province. Journal of Yunnan Agricultural University (Natural Science), 32(2), 204–211.

Chen, X. F., Tan, S. J., Liu, R. Y., Wang, Y., & Li, D. M. (2000). Study on the genetic variation of the cotton bollworm Helicoverpa armigera (Hübner) populations in China. Insect Science, 7(3), 243–249. https://doi.org/10.1111/j.1447-7197.2000.tb00415.x

Cheng, D. F., & Zhao, Z. H. (2016). Analysis of the armyworm outbreak in part areas of China and its countermeasure suggestions. Seed Science and Technology, 34(10), 89–90.

Coulon, A., Fitzpatrick, J. W., Bowman, R., Stith, B. M., Makarewich, C. A., Stenzler, L. M., & Lovette, I. J. (2008). Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma coerulescens). Molecular Ecology, 17(7), 1685–1701.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data. Proceedings of the Royal Statistical Society, 39(1), 1–38.

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(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology, 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Farrow, R. A., & McDonald, G. (1987). Migration strategies and outbreaks of noctuid pests in Australia. International Journal of Tropical Insect Science, 8(4–5), 531–542. https://doi.org/10.1017/S174275840002258X

Feng, H. Q., Zhao, X. C., Wu, X. F., Wu, B., Wu, K. M., Cheng, D. F., & Guo, Y. Y. (2008). Autumn migration of Mythimna separata (Lepidoptera: Noctuidae) over the Bohai Sea in northern China. Environmental Entomology, 37(3), 774–781.

Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate F-statistics. Journal of Heredity, 86(6), 485–486. https://doi.org/10.1093/oxfordjournals.jhered.a111627

Hao, J. H., Li, S. W., & Lin, C. S. (1992). The allozyme variation of the Mythimna separata populations in different areas. Acta Entomologica Sinica, 35(1), 33–39.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, 23(14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jiang, X. F., Jiang, Y. Y., Zhang, L., Cheng, Y. X., & Luo, L. Z. (2014b). Investigation and monitoring of overwintering and migrant populations, and the larval occurrence of the oriental armyworm, Mythimna separata (Walker). Journal of Applied Entomology, 51(6), 1114–1119.

Jiang, X. F., Luo, L. Z., & Zhang, L. (2007). Amplified fragment length polymorphism analysis of Mythimna separata (Lepidoptera: Noctuidae) geographic and melanic laboratory populations in China. Journal of Economic Entomology, 100(5), 1525–1532. https://doi.org/10.1093/jee/100.5.1525

Jiang, X. F., Luo, L. Z., Zhang, L., Sappington, T. W., & Hu, Y. (2011). Regulation of migration in Mythimna separata (Walker) in China: A review integrating environmental, physiological, hormonal, genetic and molecular factors. Environmental Entomology, 40(3), 516–533.

Jiang, X. F., Zhang, L., Chen, Y. X., & Luo, L. Z. (2014a). Novel features, occurrence trends and economic impact of the oriental armyworm, Mythimna separata (Walker) in China. Journal of Applied Entomology, 51(4), 1444–1449.

Joanna, R. F., Heather, K., & Stephen, P. (2011). Molecular ecology (2nd ed., pp. 82-147). Chichester, UK: Wiley-Blackwell.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics, 11, 94. https://doi.org/10.1186/1471-2156-11-94

Li, G. B. (1993). Reassessment overview and major progress of Mythimna separata in China. Plant Protection, 19(4), 2–4.

Li, G. B., Wang, H. X., & Hu, W. X. (1964). Route of the seasonal migration of the oriental armyworm moth in eastern China as indicated by a three year result of releasing and recapturing of marked moth. Acta Phytophylacica Sinica, 3(2), 101–110.

Li, M. M., Li, B. L., Wu, J. X., & Xu, X. L. (2018). Screening of microsatellite loci and the genetic diversity analysis of Mythimna separata (Lepidoptera: Noctuidae) populations in China. Acta Entomologica Sinica, 61(6), 712–720.

Lui, K., & Muse, S. (2005). PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics, 21(9), 1218–1219.

Luikart, G., Allendorf, F. W., Sherwin, B., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity, 89(3), 238–247. https://doi.org/10.1093/jhered/89.3.238

Lyons, J. I., Pierce, A. A., Barribeau, S. M., Sternberg, E. D., Mongue, A. J., & De Roode, J. C. (2012). Lack of genetic differentiation between monarch butterflies with divergent migration destinations. Molecular Ecology, 21(14), 3433–3444. https://doi.org/10.1111/j.1365-294X.2012.05613.x

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. Cancer Research, 27, 209–220.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America, 70(12), 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – An update. Bioinformatics, 28(19), 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Perera, O. P., & Blanco, C. A. (2011). Microsatellite variation in Helicoverpa zea (Boddie) populations in the Southern United States. Southwestern Entomologist, 36(3), 271–286.

Piry, S., Luikart, G., & Cornuet, J. M. (2001). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity, 90(4), 502–503.

Pritchard, J. K., Stephens, M., & Falush, D. (2007). Inference of population structure using multilocus genotype data: Dominant markers...
and null alleles. Molecular Ecology Notes, 7(4), 574–578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

R Core Team (2019). R: A language and environment for statistical computing. Retrieved from https://www.r-project.org/

Rosenberg, N. A. (2004). Distruct: A program for the graphical display of population structure. Molecular Ecology Notes, 4(1), 137–138. https://doi.org/10.1111/j.1471-8286.2003.00566.x

Rousset, F. (2008). Genepop’007: A complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources, 8, 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18(2), 233–234. https://doi.org/10.1038/72708

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. Science, 236(4803), 787–792.

Sun, J. T., Li, J. B., Yang, X. M., & Hong, X. Y. (2012). Development and characterization of nine polymorphic microsatellites for the small brown planthopper Laodelphax striatellus (Hemiptera: Delphacidae). Molecular Ecology Notes, 12(1), 137–138. https://doi.org/10.1111/j.1471-8286.2003.00566.x

Torriani, M. V. G., Mazzi, D., Hein, S., & Dorn, S. (2010). Structured populations of the oriental fruit moth in an agricultural ecosystem. Molecular Ecology, 19(13), 2651–2660. https://doi.org/10.1111/j.1365-294X.2010.04711.x

Wang, X. Y., Yang, X. M., Lu, B., Zhou, H. L., & Wu, K. M. (2017). Genetic variation and phylogeographic structure of the cotton aphid, Aphis gossypii, based on mitochondrial DNA and microsatellite markers. Scientific Reports, 7(1), 1920. https://doi.org/10.1038/s41598-017-02105-4

Weir, B. S. (1996). Genetic data analysis II. Sunderland, MA: Sinauer Associates.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38(6), 1358–1370.

Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. Genetics, 163(3), 1177–1191.

Wright, G. (1978). The political economy of the cotton South: Households, markets and wealth in the nineteenth century. New York, NY: WW Norton & Co.

Wright, S. (1931). Evolution in Mendelian populations. Genetics, 16(2), 97–159.

Wright, S. (1938). Size of population and breeding structure in relation to evolution. Science, 87, 430–443.

Yin, J., Feng, H. Q., Cheng, D. F., & Cao, Y. Z. (2003). Observations on the flight behavior of the oriental armyworm, Mythimna separata (Walker) in airflow. Acta Entomologica Sinica, 46(6), 732–738.

Zeng, J., Jiang, Y. Y., & Liu, J. (2013). Analysis of the armyworm outbreak in 2012 and suggestions of monitoring and forecasting. Plant Protection, 39(2), 117–121.

Zhang, G. Y., & Zhai, B. P. (2008). Construction and identification of enriched microsatellite library from Pseudaletia separata genome. Acta Ecologica Sinica, 28(8), 3860–3867.

Zhang, Y. H., Zhang, Z., Jiang, Y. Y., Zeng, J., Gao, Y. B., & Cheng, D. F. (2012). Preliminary analysis of the outbreak of the third-generation armyworm Mythimna separata in 2012 in China. Plant Protection, 38(5), 1–8.

Zhang, Y. H., Zhang, Z., Li, C., Jiang, Y. Y., & Zeng, J. (2013). Seasonal migratory behavior of Mythimna separata (Lepidoptera: Noctuidae) in Northeast China. Acta Entomologica Sinica, 56(12), 1418–1429.

Zhou, J. H., Yu, Z. H., Wu, J. K., Liu, A. Y., Hui, J. T., Zheng, Y. L., & Wu, J. X. (2017). Population dynamics of Mythimna separata (Walker) and its primary influence factors in the fall in Central Shaanxi area. China Plant Protection, 37(6), 36–40.

SUPPORTING INFORMATION

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

How to cite this article: Li M-M, Li B-L, Jiang S-X, Zhao Y-W, Xu X-L, Wu J-X. Microsatellite-based analysis of genetic structure and gene flow of Mythimna separata (Walker) (Lepidoptera: Noctuidae) in China. Ecol Evol. 2019;9:13426–13437. https://doi.org/10.1002/ece3.5799