Extensive gene flow of white-backed planthopper in the Greater Mekong Subregion as revealed by microsatellite markers

Yanqiong Yin¹, Xiangyong Li², Dong Chu², Xueqing Zhao¹, Khay Sathya³, Bounneuang Douangboupha³, Mu Mu Kyaw⁵, Manita Kongchuesin⁶, Apirusht Somrith⁶, Vinh Vien Ngo⁷, Huy Chung Nguyen⁷, Shicai Shen¹, Shufang Liu¹ & Aidong Chen¹

The white-backed planthopper (WBPH), *Sogatella furcifera* (Horváth) (Hemiptera: Delphacidae), is one of the most destructive pests in rice producing areas of the Greater Mekong Subregion (GMS) including Cambodia, Laos, Myanmar (Burma), Thailand, Vietnam, and China’s Yunnan Province. The migration sources and dispersal patterns within the region have been examined by several researchers in recent years. For example, a number of studies have begun to explore the migration sources of the WBPH based on the trajectory analyses methods¹–⁶. Due to their small body size, short lifespan, and long distance dispersal capability, it has been difficult to use fluorescent marker dyes, radar monitoring or other conventional approaches to study the migration of these insects⁷. Molecular makers have the potential to overcome those challenges⁸. Using both mitochondrial genes and microsatellite loci as molecular markers, migration patterns of the diamondback moth within China have been identified from the southern to northern regions⁹.

Our previous studies suggested that a small number of the WBPH are able to overwinter in the low latitude paddy area in the southern part of Yunnan Province, China⁴¹. The immigration sources into Yunnan Province during April to early May were estimated to be mainly from Myanmar, while the mid-May immigrations originate from northern Vietnam⁴²,⁴³. Recently, we have used the mitochondrial cytochrome oxidase I (mtCOI) marker to determine the sources and patterns of the WBPH in the GMS⁴⁴, demonstrating that the potential emigration of the WBPH from the GMS consists of three possible major routes. Those results not only confirmed that the immigration sources of the WBPH into China’s Yunnan Province were from Myanmar, Vietnam, and Laos⁴¹,⁴², but...
also indicated that Cambodia was a likely additional migration source. As nuclear genetic markers, microsatellite loci have been widely used in elucidating the genetic structure of insect populations because they are inherited codominantly and have a broad distribution and high abundance throughout the genome15–18. Determining the genetic structure of the WBPH in the GMS based on nuclear markers will provide further insights into the gene flow and migration patterns of the WBPH in this region.

In the present study, we investigated the genetic structure of 42 WBPH populations using nuclear (microsatellite loci) markers, to reveal the gene flow and migration patterns of the WBPH in the GMS. These results will benefit future sustainable management programs of this insect pest in the GMS.

Table 1. Genetic diversity indexes and genetic structure within of Sogatella furcifera collections based on microsatellite data. For each sample, the following are indicated: sampling site, population code, date of collection, host plant, sample size (N), average number of alleles per locus (Na), the effective number of alleles (Ne), the observed heterozygosity (Ho), the expected heterozygosity (He), and Nei's expected heterozygosity (Nei), the estimator of the fixation index (Fis), and the Wilcoxon test P value for heterozygosity deficit compared to expectations at mutation-drift equilibrium (Pwil). Significant values for Fis and for heterozygosity deficiency are in bold.
Genetic diversity based on microsatellite alleles. The data microsatellite locus KJ7 was eliminated due to the existence of the high null allele, thus, only eight loci were used in the analysis of genetic diversity, genetic structure, and gene flow. Values of genetic diversity indexes of the populations from China's Yunnan Province are given in Table 1. The average number of alleles per locus ($Na$) ranged from 3.375 to 8.625, and the effective number of alleles ($Ne$) ranged from 2.5234 to 5.0602. The expected heterozygosity ($He$) ranged from 0.5639 to 0.7650 while the observed heterozygosity ($Ho$) ranged from 0.2452 to 0.5563. The $He$ value in each population was higher than the value of $Ho$. Nei’s expected heterozygosity ($Nei$) ranged from 0.4922 to 0.7459. The level of genetic diversity indexes was similar in most populations.

| Collection code | IAM P-values | TPM P-values | SMM P-values |
|-----------------|--------------|--------------|--------------|
| CX              | 0.03711      | 0.42188      | 0.67969      |
| MD              | 0.32031      | 0.76953      | 0.99414      |
| BS              | 0.37109      | 0.80859      | 0.98633      |
| MS              | 0.09766      | 0.80859      | 0.98633      |
| YIJ             | 0.09766      | 0.42188      | 0.80859      |
| LC              | 0.42188      | 0.72656      | 0.97266      |
| XP              | 0.09766      | 0.67969      | 0.98633      |
| YUJ             | 0.37109      | 0.90234      | 0.98047      |
| NE              | 0.37109      | 0.96289      | 0.99609      |
| MH              | 0.47266      | 0.87500      | 0.90234      |
| SJ              | 0.15625      | 0.67969      | 0.98633      |
| GM              | 0.27344      | 0.76953      | 0.99023      |
| CY              | 0.57813      | 0.99805      | 1.00000      |
| SZ              | 0.02734      | 0.42188      | 0.97266      |
| KY              | 0.01367      | 0.80859      | 0.99414      |
| YS              | 0.23047      | 0.57813      | 0.90234      |
| FN              | 0.37109      | 0.96289      | 1.00000      |
| JP              | 0.67969      | 0.98633      | 1.00000      |
| SM              | 0.27344      | 0.90234      | 0.98633      |
| ZY              | 0.03711      | 0.42188      | 0.98633      |
| L1              | 0.37109      | 0.80859      | 0.98047      |
| L2              | 0.23047      | 0.62891      | 0.97266      |
| L3              | 0.19140      | 0.52734      | 0.97266      |
| L4              | 0.37109      | 0.80859      | 0.96289      |
| L5              | 0.52734      | 0.87500      | 0.99414      |
| L6              | 0.23047      | 0.76953      | 0.99414      |
| L7              | 0.32031      | 0.84375      | 0.98633      |
| L8              | 0.32031      | 0.84375      | 0.99023      |
| T1              | 0.46875      | 0.71094      | 0.96094      |
| T2              | 0.53125      | 0.76563      | 0.99609      |
| C1              | 0.02734      | 0.71094      | 1.00000      |
| C2              | 0.27344      | 0.80859      | 0.98633      |
| C3              | 0.01367      | 0.52734      | 0.90234      |
| C4              | 0.03711      | 0.12500      | 0.87500      |
| V1              | 0.00977      | 0.67969      | 0.99023      |
| V2              | 0.03711      | 0.52734      | 0.80859      |
| V3              | 0.28125      | 0.97656      | 1.00000      |
| V4              | 0.23047      | 0.52734      | 0.84375      |
| M1              | 0.01953      | 0.05469      | 0.65625      |
| M2              | 0.03711      | 0.37109      | 0.37109      |
| M3              | 0.02734      | 0.09766      | 0.23047      |
| M4              | 0.87500      | 0.87500      | 0.90234      |

Table 2. Within-collection tests for heterozygosity excess $P$-values according to three models (IAM, TPM, and SMM). Bold indicates significance at $P < 0.05$.

Results

Genetic diversity based on microsatellite alleles. The data microsatellite locus KJ7 was eliminated due to the existence of the high null allele, thus, only eight loci were used in the analysis of genetic diversity, genetic structure, and gene flow. Values of genetic diversity indexes of the populations from China’s Yunnan Province are given in Table 1. The average number of alleles per locus ($Na$) ranged from 3.375 to 8.625, and the effective number of alleles ($Ne$) ranged from 2.5234 to 5.0602. The expected heterozygosity ($He$) ranged from 0.5639 to 0.7650 while the observed heterozygosity ($Ho$) ranged from 0.2452 to 0.5563. The $He$ value in each population was higher than the value of $Ho$. Nei’s expected heterozygosity ($Nei$) ranged from 0.4922 to 0.7459. The level of genetic diversity indexes was similar in most populations.

The level of genetic diversity in most populations from different countries was similar. For example, the $He$ in Laos populations ranged from 0.5857 to 0.7222 which was similar to those from Thailand (0.6128–0.7113),
Cambodia (0.7279–0.7650), Myanmar (0.5639–0.6997), and Vietnam (0.6984–0.7236). The average He value in China's Yunnan populations (0.6707) was similar to those in the adjacent countries (0.6840) (P > 0.05) (Table 1).

Analyses of genetic structure within populations. The estimator of the fixation index, \( F_{is} \), was significantly different in 20 of the 42 populations, demonstrating that the presence of sub-structure within the populations was common (Table 1). In testing for deviation from mutation-drift equilibrium in BOTTLENECK, we detected a significant heterozygosity deficit (Wilcoxon test \( P < 0.05 \)) in only three populations (CY, JP, and V3). The significant heterozygosity deficit in the three populations may result from demographic expansion \(^{18,19}\) because there were no significant departures from Hardy-Weinberg equilibrium (\( F_{is} \)) in these populations (Table 1), suggesting that the significant deviation from mutation-drift equilibrium was not due to sub-structure (the Wahlund effect) within these localities.

In testing the deviation from mutation-drift equilibrium in BOTTLENECK software, we did not detect a significant heterozygosity excess in any population under the TPM or SMM models, although under the IAM model, a significant heterozygosity excess (Wilcoxon test \( P < 0.05 \)) was detected in twelve of the populations (Table 2), indicating that these twelve populations might have experienced a genetic bottleneck.

Analyses of genetic structure among populations. When considering each pairwise \( F_{st} \), 512 of 861 (59.5%) \( F_{st} \) values were associated with a significant exact test (Table 3). Analyses using STRUCTURE software identified two genetic clusters overall (\( K = 2 \)) (Fig. 1): one cluster consisted mainly of individuals from the two populations in Thailand (T1 and T2), and a few individuals from three of the populations from Cambodia (C1, C2, and C3); another cluster consisted mainly of all individuals from the other 37 populations, a few of individuals from the two populations in Thailand (T1 and T2), and most individuals from the three populations in Cambodia (C1, C2, and C3). When \( K = 3 \), the individuals from the two populations from Thailand (T1 and T2), and a few of individuals from three of the Cambodian populations (C1, C2, and C3) could also be differentiated from the other individuals.

**Table 3.** Pairwise \( F_{st} \) (genetic distance) between the collections in the Greater Mekong Subregion.

| Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation |
|---------------------|------|----------------|---------------------|------------------------|
| Among collections   | 41   | 599.456        | 0.25373Va           | 8.47                   | \( P < 0.001 \)        |
| Within collections  | 1926 | 5283.738       | 2.7437Vb            | 91.53                  |
| Total               | 1967 | 5883.194       | 2.9971              |

**Table 4.** Analysis of molecular variance (AMOVA) for structures of **Sogatella furcifera** collections.
The Mantel test results produced an $r$ value of 0.0807 for microsatellite alleles ($P = 0.8810$) (Fig. 2), indicating that no correlations were found between genetic distance and geographical distance among the populations of the WBPH in the GMS countries, indicating that extensive gene flow exists among these WBPH populations.

Significant genetic structure of the WBPH was observed at two hierarchical levels (among populations and within populations) (Table 4). Most of the variation was at the within populations level (91.53%). Although the variation among populations (8.47%, $P < 0.01$) was small, it was significant. These results demonstrated that the variations of genetic differentiation in the WBPH are mainly from inter-populations.

**Gene flow based on microsatellite data.** Based on microsatellite data, the average values of the numbers of migrants in the different countries were similar except in Thailand (Table 5). In China’s Yunnan Province, the average number of migrants was the highest in southern Yunnan, while the lowest numbers were found in the western region; this is similar to previously published results based on mitochondrial COI data. In Cambodia, Laos, Myanmar (Burma) and Vietnam, the total migrants ($N_{em}$) ranged from 883.4 (L1) to 3322.1 (M4). A high number of total migrants ($N_{em} > 1000$) were found in several populations, including L2, L4-L7 of Laos, C1, C3-C4 of Cambodia, V1-V4 in Vietnam, and M1-M4 of Myanmar, while, Thailand had the lowest number of total migrants, i.e., T1 ($N_{em} = 685.9$), T2 ($N_{em} = 807.7$). In Yunnan Province, the total migrants ($N_{em}$) ranged from 591.5 (CX) to 1224.5 (YUJ). A total of 7 populations had a high numbers of migrants ($N_{em} > 1000$), including MD in western Yunnan ($N_{em} = 1020.0$), YUJ and SM in central Yunnan ($N_{em} = 1224.5$ and $1176.0$, respectively), GM in southwestern Yunnan ($N_{em} = 1182.1$), SZ in northeastern Yunnan ($N_{em} = 1012.5$), FN in southeastern Yunnan ($N_{em} = 1207.1$), JP in southern Yunnan ($N_{em} = 1160.2$). Within the province, the number of migrants
was the highest in the southeastern region (YS and FN) (average Nem = 1071.9), and the lowest in central Yunnan (CX, SM, XP, and YUJ) (average Nem = 885.3).

**Discussion**

**Evidence for extensive gene flow in the WBPH within the GMS.** This study showed that the level of the genetic diversity in most populations originating from different countries was similar, suggesting that extensive gene flow occurs between the WBPH populations within the GMS. The heterozygosity deficit is used to test for population expansion, whereas the heterozygosity excess test is used to provide evidence of a genetic bottleneck. In this study, only three populations from China's Yunnan Province had a significant heterozygosity deficit. The results in testing the heterozygosity excess were completely inconsistent under the TPM, SMM, and IAM models. Di Rienzo et al. showed that most microsatellites fit TPM better than SMM or IAM20. Based on the TPM model, there is no significant heterozygosity excess in any of the tested populations, which suggests that no severe bottleneck effects exist in the GMS populations. Our study indicate that bottleneck effects have not played an essential role during the genetic differentiation of the WBPH. This may be due to the bottleneck effects on heterozygosity being transient and observable for only a few generations21.

**Migration sources of the WBPH within the GMS.** The extensive gene flow between WBPH populations within the GMS is consistent with our previous study9 showing that there are multiple immigration sources of the WBPH in China's Yunnan Province including Myanmar, Vietnam, Laos, and Cambodia. Although the populations from central Thailand (T1 and T2 populations) were shown to have an extensive gene flow in this study, both had limited gene flow with neighboring countries, limiting the probability of immigration of these populations into China's Yunnan Province. These results support previous results based on the trajectory analyses methods12,22,23 that the populations in central Thailand would be incapable of immigrating to Yunnan due to the lack of sufficient wind currents, incorrect wind direction and the excessive distance involved, although gene flow with the populations in Vietnam, Laos, and Cambodia would be possible23. The Cambodian populations (C1-C3 populations, especially C2) do have a limited gene flow with those from central Thailand but have extensive gene flows with other populations from other countries in the GMS. Our previous study showed that the specific mtCOI haplotype from Cambodia is only found outside the country in China's Yunnan Province (BS and GM populations). The extensive nuclear gene flow also substantiates the probable occurrence. However, the trajectory analyses methods demonstrated that the emigrant population from Cambodia would not be able to migrate to Yunnan Province23. The immigration of the WBPH individuals from Cambodia into China's Yunnan may be indirect from Vietnam or Laos, which have extensive gene flow with those in Cambodia. Whether the immigration of WBPH from Cambodia into China's Yunnan is direct or/and indirect remains to be determined.

**Future research on the source population of S. furcifera within the GMS.** The immigration sources and patterns of the WBPH as demonstrated by mitochondrial and nuclear markers are helpful in devising future, sustainable, regional management programs for this important pest in the GMS. However, the genetic basis for the migration should be further explored, including the influence of such factors as wing polyphenism24,25.
example, Xu et al. showed that two insulin receptors in the migratory brown planthopper (*Nilaparvata lugens*) (Stål) (Hemiptera: Delphacidae) play an important role in controlling long versus short wing development, providing the first evidence of a molecular basis for the regulation of wing polyphenism in insects. We had elucidated the genetic diversity of the WBPH in the GMS countries based on mtCOI and SSR markers. In future studies, it will be necessary to analyze the reliability and significance of these molecular markers relative to their consistency with WBPH biological data. With the exception of the molecular markers, the genome difference and transcriptome analyses also should be considered in a followup study. Although the 42 populations that were collected from the GMS countries help to explain the genetic diversity within somewhat limited areas, additional populations from different regions and from different seasonal occurrence in the GMS should be collected and analyzed in future studies. Additional attention needs to also be paid to more widely distributed populations, such as those from Malaysia, Indonesia, the Philippines, Bangladesh, Pakistan, India and other known occurrences of *S. furcifera* within Asia and outside of Asia to further explore and clarify the source population of *S. furcifera* in the GMS.

Table 6. Population data of *Sogatella furcifera* populations in the Greater Mekong Subregion (GMS) during 2014–2015. *all samples were collected in 2014 except those in Myanmar were collected in 2015.*
15 min, followed by 35 cycles of 30 s at 94 °C, 90 s at the primer-specific annealing temperature (Table 1), and 60 s at 72 °C, and a final elongation step at 72 °C for 30 min. The products were run on an ABI 3730xl DNA analyzer (Sangon, ShangHai, China) and the allele size was determined by comparing the mobility of the PCR products to the GeneScan™ 400HD size standard using GeneMapper software version 3.2 (Applied Biosystems, ShangHai, China).

Materials and Methods

Field sampling and DNA extraction. Adult WBPH samples were collected from 42 locations in the GMS during 2014–2015 (Table 6). The samples included 20 populations from China’s Yunnan Province, four from Vietnam, eight from Laos, two from Thailand, four from Cambodia, and four from Myanmar. The specimens were fixed in 95% ethanol and stored at −20 °C until DNA was extracted. Genomic DNA was individually extracted from each adult planthopper using the DNAzol kit (Molecular Research Center, Inc., Cincinnati, OH) and stored at −20 °C.

Microsatellite genotyping and genetic diversity based on microsatellite data. Nine pairs of fluorescent-labeled polymorphic microsatellite primers (KJ6, KJ7, KJ14, KJ16, KJ17, KJ18, KJ25, KJ35, and KJ55) (Table 7) were screened from 40 pairs of newly designed primers based on the WBPH microsatellite sequences in GenBank (until November 11, 2014) and were then used to amplify the loci using WBPH DNA as the template. The primers and the annealing temperature are described in Table 2. The PCR reactions were performed in 20 μL of a solution containing 2μL 10 × buffer, 1.5 mM MgCl2, 0.2 μM dNTPs, 1 unit Taq DNA polymerase, 2 μL template DNA, and 0.2 μM of each primer. PCR amplification was carried out as follows: initial denaturation at 94 °C for 4 min, followed by 35 cycles of 30 s at 94 °C, 90 s at the primer-specific annealing temperature (Table 1) and 60 s at 72 °C, and a final elongation step at 72 °C for 30 min. The products were run on an ABI 3730xl DNA analyzer (Sangon, ShangHai, China) and the allele size was determined by comparing the mobility of the PCR products to the GeneScan™ 400HD size standard using GeneMapper software version 3.2 (Applied Biosystems, ShangHai, China).

Based on the microsatellite alleles, the average number of alleles per locus (Na), the effective number of alleles (Ne), the observed heterozygosity (Ho), the expected heterozygosity (He), and Nei’s expected heterozygosity (Nei) of each of the 42 WBPH populations were calculated using POPGENE v.1.31. The estimator of the fixation index, Fis, was performed to detect deviation from neutrality using GENEPOP v.4.2. Wilcoxon test P value for heterozygosity deficit compared to expectations at mutation-drift equilibrium (Pwil) was calculated using ARLEQUIN v.3.5 software.

Analyses of genetic structure within populations based on microsatellite data. Deviation of the mutation-drift equilibrium in each population was tested using the BOTTLENECK software. The heterozygosity deficit was evaluated using the Wilcoxon test under the two-phase mutation model (TPM) recommended for microsatellite data. The possibility of bottleneck events within each of the 42 populations was examined under three mutation models [Two Phase Mutation Model (TPM), Infinite Allele Model (IAM), and Stepwise Mutation Model (SMM)], respectively. The TPM model was used with default settings of 30% and SMM model, 70%, respectively.

Analyses of genetic structure among populations based on microsatellite data. The traditional population differentiation approach, Weir and Cockerham’s estimator of the fixation index Fst, was calculated using GENEPOP v.3.4 software. The correlation between genetic differentiation and geographic distance was examined by Mantel test using IBDWS v.3.15 software. The distribution of genetic variation was investigated by the analysis of molecular variance (AMOVA) using ARLEQUIN v.3.5 software, and by calculating allelic diversity, heterozygosity, and pairwise values of Fst among the 42 populations. The genetic clustering of samples was examined using STRUCTURE v.2.3.2 software, using the Bayesian clustering approach with a burn-in period of 50,000 iterations and one million Markov chain Monte Carlo (MCMC) repetitions under the admixture ancestry model. Twenty independent runs were performed for each testing K value, ranging from K = 1 to 42, and ΔK was used to calculate the optimal number of genetic clusters (K).

Table 7. Sequence of microsatellite primers designed in this study.

| Code | Motif | Primer sequences | Tm (°C) | Size(bp) |
|------|-------|------------------|--------|---------|
| KJ6  | (AT)10| F:CAATGCTGCTTGTGAC R:ACAATGCACATTCA 54 298 |
| KJ7  | (CGA)5| F:CGCCGCCCTTCAATCAC T:R:AGGGTGCGTGAGAAGA 50 212 |
| KJ14 | (GTT)6| F:ATAAGCCTGACACACCCA R:AAACAGGACCAACGAGAc 54 357 |
| KJ16 | (AG)10| F:GGATTAGTGGTTGCTGCT A:R:ACCCCTGCTCTAGCATCTTT 56 271 |
| KJ17 | (TGT)7| F:CGGCGGCTTTGAAATCTAG R:AGGGTGCGTGAGAAGA 56 186 |
| AG18 | AATA35| F:ACCGGACCAGCCCTGTAAG R:GCAACGCTGGACACATT 59 212 |
| AG25 | (TG)7| F:GCGGCTACTGACAAAACAT R:CTCTCACAGGCGACTACACC 56 178 |
| AG35 | (TG)7| F:GCTTGTGTTGCGGGGCTTAG R:ACAGGGGCCTGAGGATGA 59 160 |
| AG55 | (AG)7| F:GACATTGCCCTGCTTGA R:CTGGACCACGATGAAACAT 56 127 |

Conclusions

Based on the nuclear (microsatellite) markers, the analysis of the genetic diversity, heterozygosity deficit, and heterozygosity excess suggested that there is extensive gene flow between the WBPH populations in the GMS. The genetic structure confirmed the immigration sources of the WBPH as revealed by mitochondrial markers. There is a certain gene flow between the populations in Thailand and Cambodia. It should be further explored whether the immigration of WBPH from Cambodia into China’s Yunnan Province is direct or/and indirect. These results will be helpful to the sustainable regional management of this insect pest in the GMS.

Materials and Methods

Field sampling and DNA extraction. Adult WBPH samples were collected from 42 locations in the GMS during 2014–2015 (Table 6). The samples included 20 populations from China’s Yunnan Province, four from Vietnam, eight from Laos, two from Thailand, four from Cambodia, and four from Myanmar. The specimens were fixed in 95% ethanol and stored at −20 °C until DNA was extracted. Genomic DNA was individually extracted from each adult planthopper using the DNAzol kit (Molecular Research Center, Inc., Cincinnati, OH) and stored at −20 °C.

Microsatellite genotyping and genetic diversity based on microsatellite data. Nine pairs of fluorescent-labeled polymorphic microsatellite primers (KJ6, KJ7, KJ14, KJ16, KJ17, KJ18, KJ25, KJ35, and KJ55) (Table 7) were screened from 40 pairs of newly designed primers based on the WBPH microsatellite sequences in GenBank (until November 11, 2014) and were then used to amplify the loci using WBPH DNA as the template. The primers and the annealing temperature are described in Table 2. The PCR reactions were performed in 20 μL of a solution containing 2μL 10 × buffer, 1.5 mM MgCl2, 0.2 μM dNTPs, 1 unit Taq DNA polymerase, 2 μL template DNA, and 0.2 μM of each primer. PCR amplification was carried out as follows: initial denaturation at 94 °C for 4 min, followed by 35 cycles of 30 s at 94 °C, 90 s at the primer-specific annealing temperature (Table 1) and 60 s at 72 °C, and a final elongation step at 72 °C for 30 min. The products were run on an ABI 3730xl DNA analyzer (Sangon, ShangHai, China) and the allele size was determined by comparing the mobility of the PCR products to the GeneScan™ 400HD size standard using GeneMapper software version 3.2 (Applied Biosystems, ShangHai, China).

Based on the microsatellite alleles, the average number of alleles per locus (Na), the effective number of alleles (Ne), the observed heterozygosity (Ho), the expected heterozygosity (He), and Nei’s expected heterozygosity (Nei) of each of the 42 WBPH populations were calculated using POPGENE v.1.31. The estimator of the fixation index, Fis, was performed to detect deviation from neutrality using GENEPOP v.4.2. Wilcoxon test P value for heterozygosity deficit compared to expectations at mutation-drift equilibrium (Pwil) was calculated using ARLEQUIN v.3.5 software.

Analyses of genetic structure within populations based on microsatellite data. Deviation of the mutation-drift equilibrium in each population was tested using the BOTTLENECK software. The heterozygosity deficit was evaluated using the Wilcoxon test under the two-phase mutation model (TPM) recommended for microsatellite data. The possibility of bottleneck events within each of the 42 populations was examined under three mutation models [Two Phase Mutation Model (TPM), Infinite Allele Model (IAM), and Stepwise Mutation Model (SMM)], respectively. The TPM model was used with default settings of 30% and SMM model, 70%, respectively.

Analyses of genetic structure among populations based on microsatellite data. The traditional population differentiation approach, Weir and Cockerham’s estimator of the fixation index Fst, was calculated using GENEPOP v.3.4 software. The correlation between genetic differentiation and geographic distance was examined by Mantel test using IBDWS v.3.15 software. The distribution of genetic variation was investigated by the analysis of molecular variance (AMOVA) using ARLEQUIN v.3.5 software, and by calculating allelic diversity, heterozygosity, and pairwise values of Fst among the 42 populations. The genetic clustering of samples was examined using STRUCTURE v.2.3.2 software, using the Bayesian clustering approach with a burn-in period of 50,000 iterations and one million Markov chain Monte Carlo (MCMC) repetitions under the admixture ancestry model. Twenty independent runs were performed for each testing K value, ranging from K = 1 to 42, and ΔK was used to calculate the optimal number of genetic clusters (K).
Gene flow analysis based on microsatellite data. To evaluate the dispersal of the WBPH between the populations in the GMS, the effective numbers of migrants per generation \(N_m\) was calculated using microsatellite data respectively. \(N_m\) is calculated using the mutation rate per site per generation; \(M = m/\mu\), where \(m\) is the migration rate calculated using Bayesian search strategies in MIGRATE v. 3.2.16

References

1. Huang, S. H. et al. Estimating the immigration source of rice planthoppers, Nilaparvata lugens (Stål) and Sogatella furcifera (Horváth) (Homoptera, Delphacidae), in Taiwan. Appl. Entomol. Zool. 45, 521–53, https://doi.org/10.13033/2012.521 (2010).

2. Qi, H. H. et al. Analysis of the source areas of the early immigration of white-backed planthopper in North-East Guangxi of China. Sci. Agric. Sinica 43, 3333–3342 (2011). (in Chinese).

3. Shen, H. M. et al. Analysis of the early immigration of the white-backed planthopper, Sogatella furcifera (Horváth) (Homoptera, Delphacidae) in Fujian Province, China. Acta Entomol. Sinica 54, 701–713 (2011b). (in Chinese).

4. Shen, H. M. et al. Analysis of the source areas of the early immigration of the white-back planthopper, Sogatella furcifera (Horváth) (Homoptera, Delphacidae) in Fujian Province, China. Acta Entomol. Sinica 54, 701–713 (2011b). (in Chinese).

5. Wang, Z. et al. Analysis of the source areas of early immigrant white-backed planthopper Sogatella furcifera (Horváth) in Guangdong province. Chin. J. Appl. Entomol. 48, 1253–1259 (2011). (in Chinese).

6. Hu, G. et al. Population dynamics of rice planthoppers, Nilaparvata lugens and Sogatella furcifera (Hemiptera, Delphacidae) in Central Vietnam and its effects on their spring migration to China. Bull. Entomol. Res. 107(3), 369–381, https://doi.org/10.1017/S0007485316001024 (2017).

7. Chapman, J. W. et al. Recent insights from radar studies of insect flight. Ann. Rev. Entomol. 56, 337–356 (2011).

8. Llewellyn, K. et al. Migration and genetic structure of the grain aphid (Sitobion avenue) in Britain related to climate and clonal fluctuation as revealed using microsatellites. Mol. Ecol. 12, 21–34 (2003).

9. Ma, C. et al. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. Mol. Ecol. 21, 4344–4358 (2012).

10. Wei, S. J. et al. Genetic structure and demographic history reveal migration of the diamondback moth Platella xylostella (Lepidoptera: Platellidae) from the southern to northern regions of China. PLoS ONE 8(4), e59654, https://doi.org/10.1371/journal.pone.0059654 (2013).

11. Zhao, X. Q. et al. The occurrence characteristics and population dynamics of Sogatella furcifera (Horváth) in Yunnan Province. Chin. J. Appl. Entomol. 51(2), 516–524 (2015). (in Chinese).

12. Shen, H. M. et al. Source areas and landing mechanism of early immigration of white-backed planthoppers Sogatella furcifera (Horváth) in Yunnan 2009. Acta Ecol. Sinica 31, 4350–4361 (2011c). (in Chinese).

13. Li, X. Y. et al. Possible source populations of the white-backed planthopper in the Greater Mekong Subregion revealed by mitochondrial DNAAnalysis. Sci. Rep. 6, 39167, https://doi.org/10.1038/srep39167 (2016).

14. Zheng, D. B. et al. Source areas and landing mechanisms of early immigrant population of white-backed planthoppers Sogatella furcifera (Horváth) in Shizong, Yunnan Province. Acta Ecol. Sinica 34(15), 4262–4271 (2014). (in Chinese).

15. Teng, Z. Q., Kang, L. Microsatellites reveal the genetic structure of thetoxicous strains of the migratory locust. Insect Sci. 2007, 14, 193–199.

16. Selkoe, K. A., Toonen, R. J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. Ecol. Lett. 2006, 9, 615–629.

17. Yang, X. M. et al. Invasion genetics of the western flower thrips in China: Evidence for genetic bottleneck, hybridization and bridgehead effect. PLoS One 2012, 7, e34567.

18. Cornuet, J. M. & Luikart, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001–2014 (1996).

19. Dupont, L. et al. Pathways of spread of the introduced ascidian Styela clava (Tunicata) in Northern Europe, as revealed by microsatellite markers. Biol. Invasions 12, 2707–2721 (2010).

20. Di Rienzo, A. et al. Mutational processes of simple sequence repeat loci in human populations. Proc. Natl. Acad. Sci. USA 91, 3166–3170 (1994).

21. Luikart, G. & Cornuet, J. M. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv. Biol. 12, 228–237 (1998).

22. Jiang, C. X. et al. Study on source area and landing mechanism of early immigration of white-backed planthopper (Sogatella furcifera Horváth) in Funing, Yunnan. J. Sichuan Agric. Univ. 30(2), 232–237 (2012). (in Chinese).

23. Shen, H. M. The study on overseas source area of Nilaparvata lugens (Stål) and Sogatella furcifera (Horváth) in China. Doctoral thesis, Nanjing Agricultural University (2010). (in Chinese).

24. Roff, D. A. & Fairbairn, D. J. The evolution and genetics of migration in insects. Bioscience 57(2), 155–164 (2007).

25. Xu, H. J. et al. 2015. Two insulin receptors determine alternative wing morphs in planthoppers. Nature 519, 464–467, https://doi.org/10.1038/nature14286 (2015).

26. Yeh, F. C. & et al. The user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton, Canada. (1997).

27. Raymond, M. & Rouset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumensicism. J. Hered. 86(3), 248–249 (1995).

28. Excoffier, L. & Slatkin, J. H. L. ARLEQUIN suite ver 3.5, A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10(3), 564–567 (2010).

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

30. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genome data. Genetics 155, 945–959 (2010).

31. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software structure, a simulation study. Mol. Ecol. Notes 14, 2611–2620 (2005).

32. Beerth, P. & Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. USA 98, 4563–4568 (2001).

Acknowledgements

We are very grateful to the people for kind helps during the sample population: Prof. Tao Dayun (International Cooperation Division of Yunnan Academy of Agricultural Sciences, China), Prof. Lu Li (Institute of Food Crops of Yunnan Academy of Agricultural Sciences, China), Prof. Yang Qinzhong and Prof. Zhang Fudou (Agriculture Environment and Resources Institute of Yunnan Academy of Agricultural Sciences, China). This research was supported by the Yunnan Joint Funds of the National Natural Science Foundation of China (No. U1202266), International Science & Technology Cooperation Program of China (No. 2011DFFB30040), the Yunnan Province Science and Technology Plan (No. 2014IA009), and the Yunnan Rice Industry Technology System to A. Chen,
and the Shandong Modern Agricultural Technology & Industry System (SDAIT-17-07) to D. Chu. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Author Contributions**

Conceived and designed the experiments: A.C., D.C. Performed the experiments: Y.Y., X.L., X.Z., A.C., D.C. Analyzed the data: D.C., A.C. Contributed reagents/materials/analysis tools: A.C., D.C. Wrote the paper: D.C., X.L., A.C., Y.Y. Contributed sample collection: K.S., B.D., M.K., M.K., A.S., V.N., N.C., S.S., S.L.

**Additional Information**

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

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

© The Author(s) 2017