Genetic Variation Among the Geographic Population of the Grain Aphid, *Sitobion avenae* (Hemiptera: Aphididae) in China Inferred from Mitochondrial COI Gene Sequence

XU Zhao-huan¹,², CHEN Ju-lian², CHENG Deng-fa², LIU Yong¹ and Frédéric Francis³

¹ College of Plant Protection, Shandong Agricultural University, Tai’an 271018, P.R.China
² State Key Laboratory for Biology of Plant Disease and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, P.R.China
³ Functional and Evolutionary Entomology, Gembloux Agro-Bio Tech, University of Liege, Passage des Déportés 2, Gembloux B-5030, Belgium

Abstract

In order to characterize the genetic relationship of the geographic populations of *Sitobion avenae* (Hemiptera: Aphididae) in China, a 588 bp region of the mitochondrial cytochrome oxidase subunit I (mtDNA-COI) gene was sequenced and analyzed among the different geographic populations. 269 individuals were collected from 17 localities in different wheat-growing areas in China that covered most of the range reported for this species. Within the sequence among these geographic populations, 15 polymorphic sites defined 16 distinct haplotypes, ranging in sequence divergence from 0.2% (one nucleotide) to 1.7% (10 nucleotides). Of the 15 variable sites, 12 were transitional substitutions, 2 were transversional substitutions and 1 was transitional and transversional substitution. Phylogenetic analysis showed that all haplotypes were highly interconnected with each other, in absence of phylogeographic structuring. Each of 8 haplotypes was found only at one locality, and the other haplotypes were the widespread distributed in the different localities. The higher genetic diversity was found in the northern China populations than that in the southern China populations. The low genetic differentiation ($F_{st}$=0.06945-0.69857) and high migration rate (Nm=0.21575-infinite) of Chinese populations suggest that dispersal over long distance is a major factor in the demography of *S. avenae*.

Key words: *Sitobion avenae*, insect pest, mitochondrial DNA, mtDNA-COI gene, geographic variation

INTRODUCTION

The grain aphid, *Sitobion avenae* (Fabricius) (Hemiptera: Aphididae), is an important insect pest in wheat. It probably originated in Europe, but now is widespread and occurs throughout the Mediterranean area, and in India, Nepal, China, Africa, and America (van Emden and Harrington 2007). *S. avenae* causes direct damage to crops by removing photo assimilates and acts as a vector of numerous devastating plant viruses (Dixon 1973; van Emden and Harrington 2007). In addition, the winged aphids have great capabilities of long distance migration and dispersal (Dong et al. 1987). The trait of migration over long distance of grain aphid, Barley yellow dwarf virus (BYDV) vector, would cause the virus disease epidemic in the large area, and increasing the threat to wheat production (Liu et al. 2004).

Understanding the genetic aspects of geographic variation and population structure of an insect pest can...
provide important biological information for deploying aphid-resistant cultivars, and developing chemicals alterative control strategies. Previous studies have shown weak differentiation in genetic structure of *S. avenae* population on a regional scale in Britain, France, Denmark, Romania, and Chile (Llewellyn et al. 2003, 2004; Papura et al. 2003; Figueroa et al. 2005; Jensen et al. 2008). The genetic differentiation of partial populations of *S. avenae* in China has been reported, however, such studies are limited to a small number of populations and regions (Li et al. 2001; Cai and Zhao 2004; Guo et al. 2005).

Genetic markers, in particular the sequences of mitochondrial genome, have proven to be very informative in the genetic structures and gene flow (Barrette et al. 1994; Bae et al. 2001; Cardenas et al. 2009; Liu et al. 2009; Xu et al. 2009). Because of its traits, for example, maternal inheritance, absence of intermolecular genetic recombination, a fast evolutionary rate relative to that of the nuclear DNA, the availability of efficient PCR primers (Simon et al. 1994; Hebert et al. 2004), and a wealth of comparative data (Barrette et al. 1994), mtDNA have been extensively used for studying population structures, phylogeography, and phylogenetic relationships at various taxonomic levels (Xu et al. 2009). Sequences encoding mitochondrial cytochrome oxidase subunit I (mtDNA-COI) are shown to be appropriate for intraspecific analysis due to the high degree of polymorphism observed (Hu et al. 2008).

In this study, partial sequences of mtDNA-COI gene of *S. avenae*, collected from 17 localities of major wheat-growing areas in China, were sequenced. The sequence data were used to determine the extent and character of the genetic variation of *S. avenae* populations in China.

**MATERIALS AND METHODS**

**Sample collection**

Samples of *S. avenae* geographical populations were collected from the winter wheat plant (*Triticum aestivum* L.) at 17 locations of major wheat-growing areas in China in 2009 (Table 1). Wingless aphids were sampled from different wheat plants, separated by more than 2 m in wheat field to minimize the risk of collecting the same clone, and stored at -20°C in 1.5 mL Eppendorf tubes filled with 100% ethanol prior to molecular analysis.

**DNA extraction, amplification and sequence**

Total DNA was extracted from a single individual by the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) following the protocol described by manufacturer. Approximately 700 bp fragment of the mitochondrial COI gene was amplified using the primers LepF (5'-ATTCAACCAATCATAAAGATAT TGG-3') and LepR (5'-TAAACTTCTGGATGATCC AAAAAATCA-3') (Hebert et al. 2004). Each PCR reaction was performed using a final volume of 25 μL, containing 12.5 μL 2×Taq PCR Master Mix (Biomed Biotech, Beijing, China), 2 μL 10 μmol L^-1 of each primer, and 2 μL DNA template. The thermocycling profile consisted of initial denaturation at 94°C 1 min, followed by 6 cycles of 1 min denaturation at 94°C, 1 min and 30 s annealing at 45°C, and 1 min and 15 s extension at 72°C, then followed by 36 cycles of 1 min at 94°C, 1 min and 30 s at 51°C, and 1 min and 15 s at 72°C, with a final step of 5 min extension at 72°C, and cooling to 4°C before the PCR products were removed from the thermocycler. PCR products were checked by electrophoresis on 1.5% agarose gel in TBE buffer. Amplified products were sequenced at Beijing Sunbiotech Company. Nucleotide sequences of mtDNA-COI gene were aligned using ClustalW (Thompson et al. 1994) and manually examined. Sequences were deposited in GenBank accession nos. GU138683-GU138698.

**Data analyses**

The aligned DNA sequences were compared and the variation was analysed by MEGA 4.0 (Tamura et al. 2007). For mtDNA data, the standard diversity indices, such as the number of haplotypes and polymorphic sites, haplotype diversity (h) and nucleotide diversity (π) were calculated for all sample sites with DnaSP 4.0 (Rozas et al. 2003).

The phylogenetic relationship of haplotypes was determined by Neighbor-joining (NJ) analysis (Saitou and Nei 1987), and the parsimony analysis based on the matrix of Kimura-2-parameters distances was performed using MEGA 4.0 (Tamura et al. 2007). Greenbug,
Schizaphis graminum was used as an outgroup in the study. To obtain the intraspecific phylogenies among haplotypes, the parsimony-based analysis was implemented in TCS 1.21 (Clement et al. 2000), which was used to construct a minimum spanning network of haplotypes. Because the number of samples from each population was different, the frequencies of haplotypes were not be taken into account.

Population structure analysis was conducted among S. avenae samples, genetic distance and migration rate between pairs of populations, as well as a hierarchical analysis of molecular variance (AMOVA) were estimated from mtDNA sequences in the Arlequin 3.11 program package (Excoffier et al. 2005), by calculating pairwise genetic distance ($F_{	ext{ST}}$) values and testing their significance with 1000 bootstraps. Similarly, the distances between DNA sequences were calculated by Kimura-2-parameters method (Kimura 1980). Pairwise $F_{	ext{ST}}$ values were used to estimate per-generation migration rate, Nm (the product of the effective population size $N_e$ and migration rate, m), based upon the equilibrium relationship: $F_{	ext{ST}} = 1/(2Nm + 1)$.

**RESULTS**

**MitDNA-COI gene sequence analysis**

Sequence of mtDNA-COI gene analysis of the 269 individuals of $S. avenae$, collected from 17 localities from wheat-growing areas in China, yielded 16 haplotypes (designated by H1-H16; Table 1). These haplotypes revealed 15 polymorphic sites, of which 11 were T/C transitions, one G/A transitions, one A/C transversion, one T/A transversion, and one transition and transversion (Table 2). These sequences were heavily biased toward A and T nucleotides, as expected from previous studies in aphid samples (Simon et al. 1994; von Dohlen et al. 2002). The four nucleotide acids in the sequence on averages were 41.2% T, 13.8% C, 35.4% A, and 9.6% G.

**Haplotype divergence**

Pairwise distance comparisons among $S. avenae$ haplotypes ranged from 0.2% (one nucleotide) to 1.7%
H15 (Table 3).

Eight of the 16 haplotypes were unique in their own population, 4 (H4, H6, H9, and H11) were found at 2-4 localities, indicating that most of haplotypes are locally restricted (Table 1). However, haplotypes H1, H2, H3, and H5 were shared by 8-14 populations, respectively (Table 1). Collectively, the distribution can be characterized by the co-existence of main locally restricted and minor widely distributed haplotypes.

Phylogenetic and network analyses

Phylogenetic relationships among haplotypes in the study and the documented sequences of *S. avenae* are depicted in Fig. 1. Most of haplotypes were weakly associated (less than 50% bootstrap support) or unresolved possibly due to small nucleotide acid difference among them (Table 2). Haplotypes (H3, H13, and H15) and documented *S. avenae* clones (EU701907 and GU667465) obtained marginal support as a group (Fig. 1). But there was no evidence for strong geographical clustering in the trees.

To further illustrate the genetic relationships among *S. avenae* haplotypes, a minimum spanning network, which visualizes a possible evolutionary pathway among closely related haplotypes, was obtained (Fig. 2), but
the network analysis provided very limited information. All haplotypes were highly interconnected with each other, it seemed that no haplotype or haplotype group had diverged.

Population genetic structure

Table 1 shows that the nucleotide diversity (\(\pi\)) and the haplotype diversity (h) within each population ranged from 0.00276 to 0.00875 and from 0.325 to 0.867, respectively. Samples from YH possessed the lowest nucleotide diversity (0.00276±0.00106) and haplotype diversity (0.325±0.125), while samples from STA and HLA had the highest nucleotide diversity (0.00814±0.00092) and haplotype diversity (0.867±0.048), respectively. The results showed that the genetic diversity of S. avenae in northern areas was higher than that in southern areas, and haplotype diversity of most of localities in the north were higher than 0.700.

Data from AMOVA molecular detection proved that there was 18.91% genetic variation in inter-populations (Table 4), which illustrated that genetic differentiation of S. avenae was mainly occurred among the groups of inter-population. Genetic distance (\(F_{ST}\)) and per-generation migration rate (Nm) between pairs of 17 populations are shown in Table 5. Pairwise genetic distance (\(F_{ST}\)) among 136 pairs of populations ranged from -0.06945 to 0.69857. Among them, 66 showed statistically no significant genetic differentiation (\(P>0.05\)), suggesting that approximately 50% pairs of populations form one genetic group (Table 5). These results reveal the lack of genetic structure in S. avenae among the sampled areas, and agree with the network analysis.

In addition, gene flow among the 17 localities was estimated by Nm, which is the expected number of migrants exchanged among populations in each generation. According to the Nm values between pairs of populations (Table 5), YH with other 16 populations, XS with other 13 populations (except HLA, HS and SB), and SB with HZ, were all less than 1, while other pairwise Nm values were greater than 1. These results

| Table 4 Analysis of molecular variance (AMOVA) of S. avenae populations |
|------------------------|------------------|---------------|--------|------|
| Source of variation    | df   | Variance components | Percentage (%) | P    |
| Inter-populations      | 16   | 0.45167*            | 18.91          | <0.001|
| Within populations     | 252  | 1.93629*            | 81.09          | <0.001|
| Total                  | 268  | 2.38795             |                 |      |

*, level of significance at P<0.001.
suggest that the extensive gene flow occur among *S. avenae* populations in China.

**DISCUSSION**

Generally aphids have low mtDNA divergence. Divergence of only 0.4% was found in a study using the mtDNA-COI gene, cited as evidence against the hypothesis that there were host races in the pea aphid, *Acyrthosiphon pisum* (Boulding 1998). In this study, the maximum mtDNA sequence divergence in *S. avenae* was 1.7%. *S. miscanthi* (Takahashi) and *S. avenae* differ by only 1.5% sequence divergence in the mtDNA-COI gene (Sunnucks and Hales 1996). Thus, the magnitude of sequence divergence in *S. avenae* is comparable to that revealed in similar studies.

Phylogenetic tree (NJ) and the minimum spanning network of the 16 haplotypes suggested that most haplotypes were weakly associated (less than 50% bootstrap support) or unresolved. A similar result is also reported for other insect pests (Li et al. 2006; Hu et al. 2008). These results indicate a closely phylogenetic relationship among *S. avenae* haplotypes. Further, haplotypes H1, H2, H3 and H5 were shared by 8-14 populations, respectively (Table 1). Based on the geographic distances, the occurrence of identical haplotypes over such a wide areas is noteworthy.

In this study, the results showed that the genetic diversity of *S. avenae* was higher in northern areas than that in southern areas, and haplotype diversity was higher than 0.700 in most of the northern localities, which deduced that *S. avenae* in northern sampled areas were immigration from different distance of southern area, because the aphid is unable to survive through the winter in north of China, but could overwinter in the south and migrate into the north in spring. *S. avenae* individuals with different haplotypes immigrate, reproduce through parthenogenesis and then damage to host plants during the wheat growing season, which increase the genetic diversity of northern populations.

The sampling sites covered a wide range of major wheat-growing areas (17 localities). For example, the straight line distance between JZ and QX is approximately 1690 km. However, all populations, except from YH, SJ, XS, and SB, hardly showed any differences in genetic distance (*F**st*) and per-generation migration rate
(Nm). Generally, if Nm<1, local populations will develop differentiation; if Nm>1, there will be little differentiation among populations (Wright 1951). According to the Nm values, the pairs of YH with other 16 populations, XS with other 13 populations (except HLA, HS and SB), and SB with HZ were all less than 1, while other pairwise comparisons of Nm values were greater than 1. These results suggest that gene flow among S. avenae populations in China may prevent natural population from genetically diverging by genetic drift. The long distance migration of the grain aphid may enhance gene flow, which is consistent with the previous studies (Close and Tomlinson 1975; Dong et al. 1987; Luo et al. 1988; Yang 1990; Simon et al. 1999; Li et al. 2001; Cai and Zhao 2004; Llewellyn et al. 2004; Guo et al. 2005). Such genetic structure is a typical trait of migratory insect.

S. avenae is a migratory pest insect, which is widely distributed throughout the wheat growing regions of China. We analyze partial sequences of the mtDNA-CoI gene of S. avenae to determine the extent and nature of their genetic variation in China. The results present evidence to support the previous studies of the migration capability, and provide an important theoretical basis for deployment of aphid-resistant wheat in different wheat-growing regions.

Acknowledgements

This study was supported by the National Basic Research Program of China (973 Program, 2006CB102004), the National S&T Major Project of China (2009ZX08012-007B), the National Natural Science Foundation of China (30971920), and the Inter-University Targeted Project between Belgium and China (PIC SHANDONG).

References


(Managing editor ZHANG Juan)