Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, *Laodelphax striatellus*

Jing-Tao Sun¹, Man-Man Wang¹, Yan-Kai Zhang¹, Marie-Pierre Chapuis², Xin-Yu Jiang¹, Gao Hu¹, Xian-Ming Yang¹, Cheng Ge¹, Xiao-Feng Xue¹ & Xiao-Yue Hong¹

¹Department of Entomology, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China, ²CIRAD, UMR CBGP, F-34398 Montpellier, France.

Dispersal not only has important consequences for the spatial and temporal dynamics of populations, but also affects gene frequencies and the relative importance of a given species in the local community. Dispersal also helps an organism to cope with environmental changes by allowing it to colonize more suitable habitats. Besides its ecological importance, dispersal has effects on evolutionary processes. For example, speciation depends on a balance between local adaptation and dispersal. Consequently, measuring dispersal is a focus of research in evolutionary biology, conservation biology and pest management.

For pest species, quantifying dispersal is not only key to resolving questions related to pests’ behavioral traits, and physiological or genetic constraints, but also to developing a forecasting system for integrated pest management plans. A case in hand is the small brown planthopper (SBPH), *Laodelphax striatellus* (Fallén), a notorious agricultural pest, which is widely distributed from the Philippines to Siberia and Europe, mainly in the temperate zone. SBPH not only causes sucking damage to crops but also spreads the damaging plant disease *Rice stripe virus* (RSV). The most recent outbreak of RSV in eastern China in 2000 caused significant losses. RSV is strictly transmitted by SBPH in persistent propagative and transovarial manners. Therefore, the dispersal ability of SBPH plays a critical role in shaping the population structure of RSV. Sometimes, it is unclear whether the outbreak of RSV was triggered by local strains or invasive strains associated with the dispersal of SBPH. Understanding the dispersal pattern of SBPH is therefore not only key to developing effective control strategies for this pest, but also has implications for the ecology and evolution of vector-borne RSV virus systems, as well as for RSV epidemiology and control.

Two alternative views exist on the importance of dispersal in SBPH population dynamics. The common view is that outbreaks are caused by local resident populations. Unlike the other two rice planthoppers, brown planthoppers (*Nilaparvata lugens*) and white-backed planthoppers (*Sogatella furcifera*), SBPH is able to overwinter by diapausing as third- or fifth-instar nymphs in temperate climates, including all regions of China, and no seasonal migration is required between climatic areas of the species range. Moreover, SBPH populations display geographic differences in the critical day length to induce larval diapause and in the diapause depth. If genetically-induced, such life history variation among SBPH populations would suggest local adaptation to climate and
limited gene flow between populations. This is supported by a population genetic structure of SBPH in Japan based on three allozyme loci which reveal significant differentiation between populations.

An alternative view is that outbreaks are caused by long-distance migration events. Potential to migrate is indicated by catches of a considerable number of SBPH at high altitudes or over the sea with *N. lugens* and *S. furcifera*, two species which annually migrate to China from the Indochina Peninsula riding the air currents in the early spring. Moreover, a comparison of insecticide susceptibility between Chinese and Japanese populations coupled with a backward trajectory analysis indicated that SBPH can fly from China to Japan. Accordingly, extensive gene flow between populations was suggested by a phylogenetic analysis in China based on a fragment of the Cytochrome C Oxidase Subunit II (*COII*), which showed that identical or highly similar haplotypes were shared between different geographic populations. These observations suggest that SBPH, like *N. lugens* and *S. furcifera*, is able to migrate long distances.

Conventional methods for studying the movement of insects include fluorescent marker dyes and radio-isotopes. However, these methods are difficult to apply to SBPH because of their small size, short lifespan, large population sizes, and rapid aerial population migration. Alpha-proteobacteria and induces cytoplasmic incompatibility (CI) in the range of insect species. However, there is not yet any conclusive evidence for the population genetic structure of SBPH. The molecular markers (allozymes or *COII* gene), numbers of samples (9–281) and sample types (laboratory reared populations or natural populations) used in previous studies have led to contradictory results and limited ability to reveal the reliable pattern of its population structure.

Furthermore, by its influence in shaping host population genetics, infection by the endosymbiotic bacterium *Wolbachia*, which is common in SBPH, may at least partially explain the contradictory results of previous studies. *Wolbachia* is a maternally inherited alpha-proteobacteria and induces cytoplasmic incompatibility (CI) in SBPH, a mechanism by which mating between uninfected females and infected males results in no offspring. However, the infected females can mate with both the infected and uninfected males, and produce offsprings successfully. The CI level of SBPH is quite high and even old males strongly cause CI. Because of the reproductive advantage of *Wolbachia*-infected females, even a small number of *Wolbachia*-infected insects can lead to a rapid spread and fixation of the infection. In association with *Wolbachia* spread, a small number of mitochondrial haplotypes borne by infected individuals are also spread within populations, reducing the mitochondrial diversity of populations. Moreover, CI can prevent gene flow from infected to uninfected populations, promoting genetic divergence, and has attracted attention as a possible mechanism for rapid speciation. Thus, the high CI level in SBPH can bias measurements of diversity and structure estimated by mtDNA markers in its populations. However, if complemented with data from nuclear markers, the effect of CI on the population structure of its host can be assessed and controlled. Simultaneously, the *Wolbachia* infection status (infection frequencies and *Wolbachia* strains) of different SBPH geographic populations per se will provide information on the population structure and dispersal ability.

In this study, we set out to explore the following two questions: (1) Does SBPH show evidence of a high level of dispersal from population genetic data? If this is the case, we expect a low level of genetic differentiation between populations and no isolation-by-distance effect. (2) Does *Wolbachia* affect the mitochondrial DNA genetic diversity and structure of SBPH populations? If this is the case, we expect a low level of genetic diversity at mitochondrial markers and congruent geographic structures of *Wolbachia* and mitochondrial genomes. To address these questions, we evaluated the spatial genetic variation in 26 SBPH populations sampled from most of the species’ range in China. To gain a reliable picture of genetic variation, we employed three different types of molecular markers: (i) 13 microsatellite loci that had previously been published (6) or were presently identified from an expressed library (7); (ii) two mitochondrial DNA gene sequences (*COI* and *COII*); and (iii) the multilocus sequence typing of Baldo et al. (2006) for the endosymbiotic bacterium *Wolbachia*.

**Results**

**Genetic diversity.** The concatenated mitochondrial sequences of the *COI* and *COII* genes were obtained from 1323 out of 1328 samples. In the concatenated sequence (1368 bp), 85 of the nucleotides (6.2%), were variable between the 1323 SBPH samples. Of the 85 variable sites, 39 were parsimony informative sites. No nuclear mitochondrial DNA sequences (numts) were identified after comparisons with the published complete mitochondrial genome and examinations of sequence translation. Of the 1323 samples, 107 haplotypes were identified (Fig. 1B), of which 70 were unique haplotypes, five were shared within populations and 32 were shared between populations. Three haplotypes of high frequency accounted for 73.5% of the haplotypes. H1 was the most common haplotype, being present in 620 individuals distributed in all 26 populations. H10 was found in 225 individuals from 23 populations and H2 was found in 128 individuals from 24 populations.

Table 1 shows genetic diversity indices for each SBPH population sample. The populations showed moderately high haplotype diversity and low nucleotide diversity at mtDNA sequences. Overall haplotype diversity and nucleotide diversity were 0.739 and 0.18%, respectively. At microsatellite markers too, the SBPH populations showed a high level of genetic diversity. The allelic richness (*A_R*) based on a minimum population size of 23 diploid individuals ranged from 8.3 in DD10 to 10.1 in SH with an average value of 9.1 across populations. The unbiased expected heterozygosity (*uH_E*) values were quite high, ranging from 0.752 to 0.810 (Table 1, Supplementary Table S1 online). Details of single-locus indices of genetic diversity can be found in Supplementary Table S1.

Figure 2 shows the results of a comparison of genetic diversity between climatic zones. For the mitochondrial markers, both the levels of haplotype diversity (*H_D*) and the levels of nucleotide diversity (*π*) were significantly lower in the moderate temperate zone (MT) than in the three remaining climatic zones: WT (warm temperate zone), SS (southern subtropical zone) and NS (northern subtropical zone). For *H_D*, the *P* values were <0.001 for MT vs. WT and MT vs. NS, and = 0.03 for MT vs. SS. For *π*, the *P* values were <0.001 for MT vs. WT and MT vs. NS and = 0.014 for MT vs. SS. In addition, the levels of nucleotide diversity (*π*) were also significantly lower in SS than in WT (*P* = 0.048) and NS (*P* = 0.020). For the microsatellite markers, the only significance indicated that the levels of *A_R* were significantly higher in MT than in WT (*P* = 0.024).

**Population structure.** Pairwise *F_{ST}* values computed from mitochondrial DNA data ranged from −0.078 to 0.442, with an average *F_{ST}* of 0.075. Exact tests showed that 120 of the 325 pairwise populations were significantly different (Fig. 3A and Supplementary Table S2 online). Interestingly, 90 of the 120 significant pairwise differentiation levels were between populations from the MT zone and the populations from the three remaining climatic zones. The average *F_{ST}* value between populations from the MT zone and
populations from the three remaining climatic zones was 0.176. In comparison, the F_{ST} values for other between-zone population pairs were very low, with an average F_{ST} of 0.016. The principal coordinate analysis (PCoA) also showed that the six populations from the MT zone (MDJ, HRB10, HRB12, SJ, YJ11 and YJ12) were separated from the other populations (Fig. 4A). A Bayesian inference (BI) phylogenetic tree and the haplotype network derived from 107 concatenated mtDNA sequences revealed two well-supported haplogroups (Fig. 1A and 1B). The mean P-distance between the two haplogroups was 0.4%. These haplogroups, named HGI and HGII, had non-random distributions across the sampling localities. In general, the frequency of HGI (61.3%) was higher than that of HGII (38.7%) over all the sampled populations, and had absolute dominance in the MT zone (95.1% in total, Fig. 5; 94.6% in females, and 97.9% in males, Supplementary Fig. S1A and 1B online).

Pairwise F_{ST} values computed over microsatellite loci were very low, ranging from −0.017 to 0.027, with an average F_{ST} of 0.006 (Fig. 3B and Supplementary Table S2 online). The standardized pairwise F_{ST} values ranged from −0.082 to 0.114, with an average F_{ST} of 0.027 (Fig. 3C and Supplementary Table S3 online). We did not
detect significant genotypic differentiation between any pairwise populations by using exact tests. For the principal coordinate analysis (PCoA), all of the populations except TS showed little genetic structure (Fig. 4B), though the pairwise $F_{ST}$ values between TS and the other populations were very low (0.002–0.027, Fig. 3B, Supplementary Table S2 online). Accordingly, STRUCTURE analyses suggested that SBPH most likely forms a single genetic cluster. Indeed, for $K = 1$, the log-likelihood of the multilocus genotypic data was maximal and had low variance (see Supplementary Fig. S2 online).

### Geographic and environmental associations

In our microsatellite data set, no isolation by distance (IBD) effects were detected with the standardized pairwise $F_{ST}$ ($Z = 17.184, r = -0.071, P = 0.735$; Fig. 6A). By contrast, there was a significant IBD effect across the 22 geographic populations in the mtDNA sequence data ($Z = 65.357, r = 0.383, P = 0.001$; Fig. 6B). When the populations from the MT zone were excluded, we were no longer able to detect the IBD effect across the remaining 16 populations ($Z = 7.794, r = -0.093, P = 0.778$).

The logistic regression analyses by SAM software indicated that the distributions of the two mtDNA haplogroups were significantly associated with the first two PCA axes ($P < 0.001$). PC1 and PC2 axes accounted for 71.1% and 13.4% of the climatic variation across the sampling sites. Supplementary Table S4 shows that PC1 was primarily related to the mean (or maximum) of temperatures while PC2 mostly explained the amount of precipitation and climate variability (seasonality of precipitation and isothermality). The values of mean and (maximum) temperatures, isothermality, annual precipitation, seasonality of precipitation and annual relative humidity for each sampling location are presented in Supplementary Table S5. Further, partial Mantel tests showed that HGI frequencies were significantly correlated with PC1 ($r = 0.437, P < 0.001$) but not with PC2 ($r = -0.055, P = 0.300$). When the IBD effect was removed, the significant correlation between the HGI frequencies and PC1 was still detected ($r = 0.237, P < 0.05$).

### Demographic history

For neutrality tests, significantly negative Tajima’s $D$ values ($-2.127, P < 0.001$) and Fu’s $F_{S}$ ($-26.018, P < 0.001$) were found over all populations. The mismatch distribution over all populations was deemed unimodal and failed to reject the hypothesis of a sudden expansion because of the small, non-significant Harpending’s raggedness (HR) index (0.063, $P = 0.294$, 4).
SCIENTIFIC REPORTS | 5 : 8045 | DOI: 10.1038/srep08045

**Discussion**

**The effect of Wolbachia on host evolution.** The cotransmission of Wolbachia and mitochondria is expected to cause the bacteria to have an indirect impact on mitochondrial DNA diversity, as a result of a selective sweep of the mito-type associated with the infection. In this study, all individuals were infected with the same or at least very closely related Wolbachia strains. This compares to an average infection frequency of 91.7% reported for SBPH populations in China and Vietnam and an average infection frequency of 65.9% for SBPH populations in Japan. The rapid growth in the infection frequencies of the same or at least very closely related Wolbachia strains indicates a recent invasion of this endosymbiont in SBPH. Since these strains can induce strong cytoplasmic incompatibility (CI) in SBPH, limited mtDNA diversity would have been expected. However, haplotype diversity was moderately high (0.739), suggesting no or little effect of Wolbachia on mtDNA diversity. The observation that different mtDNA haplotypes carry the same or very closely-related Wolbachia strains strongly indicates horizontal transmissions of Wolbachia between SBPH individuals. A similar phenomenon has also been reported in the fire ant Solenopsis invicta and the fruit fly Drosophila willistoni. The horizontal transmissions of Wolbachia make the natural infection status unstable in SBPH populations, and also interrupt the host speciation process induced by CI. Our data suggest that Wolbachia alone may not be sufficient to promote population divergence and speciation of the host over long periods of evolutionary time.

**Dispersal pattern of SBPH.** In agreement with Ji et al. (2010), the SBPH populations in China were found to show a high degree of genetic connectivity, despite a high level of polymorphism at the two types of molecular markers. This population structure pattern is consistent with those of other migratory insects that are characterized by genetic homogeneity between populations separated by large geographic distances. The population genetic structure of SBPH in China contrasts with the strong population genetic structure of other insects that are widely distributed, but which have limited dispersal ability, such as the striped rice borer Chilo suppressalis (Walker).

Several lines of evidence support the high dispersal ability of SBPH. First, our microsatellite results from pairwise $F_{ST}$, PCoA, and Bayesian clustering analyses indicated a lack of geographic structure among populations. The PCoA result of mitochondrial DNA also showed that all of the populations were similar except those from the MT zone (i.e., MDJ, HRB10, HRB12, SJ, YJ11, and YJ12; and see below). In addition, there was no obvious evidence for isolation by distance (IBD). Usually, IBD effects are pronounced in moderately mobile species but weak in both low-mobility and high-mobility
species, since extensive gene flow homogenizes populations across large geographic areas. Furthermore, the SBPH populations in China possess high levels of mtDNA haplotype diversity with low levels of nucleotide diversity. This diversity combination has been considered to be mostly the result of demographic expansion after a period of small effective population size, retaining new mutations. SBPH populations appear to have experienced a recent demographic expansion, and this conclusion was supported by both mismatch and BSP analyses in our study. However, this diversity combination type may also reflect a high level of dispersal in SBPH since it allows for the sweeping away of deleterious mutations that had built up within isolated populations. The low level of structure observed among populations suggests that outbreak occurrences of SBPH and RSV may not only depend on local recruitment but also on immigration. However, the lack of genetic differentiation prevented us from quantifying migration rates between populations and the frequency of long-distance migration events. Although assignment tests based on multilocus genotype data can help to estimate recent migrations, the accuracy largely depends on the degree of genetic differentiation between populations. Computer simulations have also shown that as the dispersal rate increases, and genetic differentiation decreases, the ability to distinguish the source of an individual also falls. The lack of population differentiation between SBPH populations makes this analysis impractical.

**Discordance between mitochondrial DNA and nuclear microsatellites.** The mtDNA haplotypes had a non-random distribution across the sampling localities in the face of gene flow (Fig. 5 and Supplementary Fig. S1 online). The levels of mitochondrial diversity were significantly lower in the MT zone than in other climatic zones. Conversely, the level of microsatellite diversity was generally similar across climatic zones (Fig. 2). Mitochondrial HGI had a high frequency in the MT zone, and had temporal stability (Fig. 5 and Supplementary Fig. S1 online). The PCoA result for mtDNA also showed that the six population samples from the MT

![Figure 5](imageurl) | Sampling locality and the distributions of mitochondrial DNA haplogroups of the SBPH populations. CT, cold temperate zone; MT, moderate temperate zone; WT, warm temperate zone; NS, northern subtropical zone; SS, southern subtropical zone. MaTr, marginal tropical zone; MTr, middle tropical zone; ARDC, ancient rice domestication center; HGI, mitochondrial DNA haplogroup I; HGII, mitochondrial DNA haplogroup II (The map is made by ArcGIS 10.2 software, http://www.arcgis.com/features/).

![Figure 6](imageurl) | Scatter plots of genetic distance ($F_{ST}/1-F_{ST}$) vs. geographic distance for pairwise population comparisons based on microsatellite (A) or mtDNA data (B). Blue circles represent the comparisons associated with populations from moderate temperate zone (MT).
zone (MDJ), HRB10, HRB12, SJ, YJ11 and YJ12) were genetically different from the other populations. However, the microsatellite data revealed a lack of genetic structure in China (Fig. 4). The lower differentiation levels at microsatellites remained even after standardization (Fig. 3C, Supplementary Table S3 online). So why is there such a large haplotype skew over nuclear data in the MT zone for SBPH? The biogeography of mitochondrial and nuclear discordance in animals has been well reviewed and meta-analyzed by Toews and Brelsford (2012)44. Mito-nuclear discordance arising in the face of strong gene flow has rarely been reported. Usually, mitochondrial introgression from a close species, male-biased dispersal, demographic expansion or selection on mtDNA are considered as the main reasons for this mito-nuclear discordance45–49.

In our data and literature, there was no evidence for hybridization of SBPH with a recently formed species or divergent population, and the divergence between the two mitochondrial haplogroups was low (P-distance = 0.4%). There was no evidence for a contribution of male-biased dispersal to mito-nuclear discordance either. Male-biased dispersal is common in nature, and unlike males, macropterous females of SBPH can decrease in number with low population density26. This suggests higher or longer-distance dispersal in males, at least when the population density is low, but under such a hypothesis, the prevalence of HGII should be expected to be higher in the male than in the female populations from the MT zone. Similar patterns of distribution of the two haplogroups for both genders do not support the male-biased dispersal hypothesis in this species. This is probably explained by the same high rate and long distance of dispersal for females as for males when the population density is high. When population sizes are large, the resulting large number of migrants prevails in affecting population genetic differentiation.

An alternative scenario would be that SBPH has experienced a recent northward colonization associated with genetic drift effects. Structuring effects in the northeastern region (i.e. MT zone) would be expected to have a more pronounced impact in the mitochondrial genome than the nuclear genome, since mtDNA is haploid and maternally inherited, and therefore has a fourfold smaller effective population size. The lower mtDNA diversity and significant differentiation in the MT zone are congruent with this scenario. The reason for northward expansion is unclear here. However, some consideration should be given to the history of rice domestication since, although polyphagous, SBPH mainly occurs on this host plant. According to archaeological evidence, the eastern region (NS zone) of China is one of the rice domestication centers (Fig. 5)51, and, 5700–3000 years ago, the cultivation area expanded in the northern and southern regions of China52. Rice cultivation in the northeastern region (MT zone) started more recently, around the end of the 19th century. The expansion of rice growing areas may have provided suitable food sources for the SBPH populations, and subsequently triggered a range expansion northward and southward from the eastern center.

A final alternative explanation is local adaptation of mitochondrial haplogroups to different climates. The finding that the mitochondrial HGI frequencies were significantly correlated with the first PCA axis (even when the IBD effect was removed) may indicate that temperature plays an important role in shaping the distributions of mitochondrial haplogroups. Considering the low winter temperatures in the MT zone (the mean winter temperatures of the four sampling sites in the MT zone range from −17.1 to −12.0°C), the enrichment of HGI in the MT zone may indicate that the HGI mitochondrial haplotype (or another allele at another locus it is linked with) is more favored under cold climates than the HGII mitochondrial haplotype. Mitochondria not only provide most of the energy (ATP) in animal cells through oxidative phosphorylation (OXPHOS), but also provide heat to maintain body temperature through proton leakage53. Certain mitochondrial lineages (or haplogroups) characterized by reduced coupling efficiency of OXPHOS and increased efficiency of proton leakage have been suggested to adapt to colder climates, resulting in their enrichment in such climates54–56. Although we did not find any non-synonymous mutations at the two studied cytochrome oxidase genes, we cannot rule out the probability that their genetic variation is impacted by selection acting on the remaining 11 mitochondrial protein-coding genes through genetic hitchhiking. Further studies are needed to identify the candidate genes for cold adaptation by whole-genome comparisons and functional assays of the two types of mitochondria.

The traditional view that sequence variation in the mtDNA genome is selectively neutral is being challenged by a growing body of evidence for natural selection on mitochondrial genes54–56. Recently, Kazancıoğlu & Arnqvist also documented negative frequency-dependent selection (NFDS) on mtDNA haplotypes in laboratory populations of the seed beetle (Callosobruchus maculatus)57. Whether the maintenance of the mtDNA diversity and the distribution of the mtDNA variations in SBPH populations are associated with NFDS, or the combined effect of NFDS and climate adaptation need to be studied in the future.

Methods
Sample collection and DNA extraction. The study area consisted of 22 sampling locations along a gradient covering four climatic zones (moderate temperate zone, MT; warm-temperate zone, WT; northern tropical zone, NT; and southern tropical zone, ST) in China58. SBPH adults were sampled from rice plants at 22 sites during the summers (May to August) of 2010, 2011 and 2012. Four sites (Harbin, Yanji, Dandong and Hefei) were sampled twice in two years (Table 1, Fig. 5). The

Figure 7 | Mismatch distributions (A) and Bayesian skyline plots (B). (B) Dark lines represent mean inferred effective population size (Ne) multiplied by the generation time (T) in years, blue areas mark the 95% highest probability density (HPD) intervals. Ne × T is presented on a logarithmic scale.
geographic distances between sampling sites varied from 138 to 3071 km. To avoid samplingsiblings, we collected only one SBPH per host plant and selected host plants that were at least 1 m apart. The samples were immediately preserved in absolute ethanol and then stored at −20 °C in the laboratory until DNA extraction. In all, 1140 adult females were collected, with sample sizes ranging from 6 to 48 per population. In contrast, only 188 adult males were collected from 15 of the 22 sites (Table 1). Total genomic DNA was isolated from the combined head and thorax of each individual using a modified CTAB method.

Climate GIS layers for the study area were downloaded from the online WORLDCLIM database (www.worldclim.org). The BIOCLIM algorithm implemented in DIVA-GIS software was used to derive 19 climatic variables for 21 of the 22 sampling sites (except for NN) from three input parameters (mean monthly values of maximum temperature, minimum temperature, and precipitation). Because many of these climatic variables were likely to be correlated, we used SPSS 17.0 to perform a principal component analysis (PCA) on a correlation matrix of the 19 climatic variables.

Mitochondrial DNA sequencing. Portions of the COI gene (895 bp) and COII gene (650 bp) of mtDNA were separated by PCR for all samples. The primer pair for the COI gene (F: 5'-TCTCATTATACATCATCAGTTAGTTAG-3'; R: 5'-GTAGTCCTGAATATCGTGGTATTTGATT-3') was designed from the published complete mitochondrial genome of S. tabaci using Primer premier 5.0 software (http://www.premierbiosoft.com). The COII gene fragment was amplified with primers COII-1 and COII-2. PCR was performed in a Veriti Thermal Cycler (ABI Biosystems). All PCR reactions consisted of 40 μl reaction volumes containing 10–100 ng template DNA, 0.2 μl of DreamTaq polymerase (5 U/μl), 4.0 μl 10× DreamTaq Buffer (including 20 mM MgCl2), 4 μl dNTPs (2.0 mM each), 1 μl primers (25 μM) and autoclaved distilled water to volume. The thermal profile used an initial denaturation step of 95 °C for 3 min followed by 35 cycles of denaturing at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 1 min. A 15 min final extension at 72 °C was added at the end of the cycle. Negative controls were included in both DNA isolation and PCR reactions. After verification by gel electrophoresis, the PCR templates were purified and then sequenced in both directions using the same primer pairs on an Applied Biosystems 3130 Genetic Analyzer.

Microsatellite development, genotyping and null alleles. New microsatellites were isolated from the expressed sequence tag of SBPH, which were downloaded from the Short Read Archive of the National Center for Biotechnology Information (NCBI) with accession numbers SRX016333 and SRX016334. Data mining and experimental procedures were done according to Sun et al. (2011)87. Seven new microsatellites were developed and proved to be highly polymorphic, with 7 to 28 alleles per locus (see Supplementary Table S1 for primer sequences and allele profiles). These new microsatellites, together with six previously developed microsatellite loci developed by Sun et al. (2012)88, were genotyped. Each of the 12 loci was amplified in a single tube separately for each individual. PCR for the seven newly developed microsatellites was performed according to Sun et al. (2011)87. PCR for the six previously developed microsatellite loci was performed according to Sun et al. (2012)88. Products were genotyped on an Applied Biosystems 3130 Genetic Analyzer.

No linkage disequilibrium between loci was detected. A large excess of significant departures from HWE was observed in 130 of the 338 single-locus exact tests after false discovery rate (FDR) correction94. Two loci (LS1, LS3, LS4, LS5 and LS6) accounted for 97.6% of the significant outcomes. Significant deviations from HWE were revealed by the Checkr software (Ver. 2.2.3)95. The deviation from HWE was almost entirely due to heterozygote deficiency. Only two cases (locus LS1 in J and LS4) were called by heterozygote excess (see Supplementary Table S1 for more details). As null alleles may inflate genetic differentiation96, all of the population structure analyses below were conducted separately on two data sets: the original data set of 13 loci and a reduced data set of eight loci (i.e., excluding LS1 and LS3–LS6 with high null allele frequencies). Since highly congruent results were obtained in these analyses, we only present the results obtained with the original data set.

Screening for Wolbachia and multilocus sequence typing (MLST). We first checked for the presence of Wolbachia infection in all 1328 samples, by using PCR primers wsp81F and wsp819R89. PCR was performed according to Zhang et al. (2013)90 and 13 individuals were tested twice to assess repeatability. All amplifications, including positive and negative controls, were checked on 1% agarose gel. We then selected all males and 12 Wolbachia-positive females from each population sampled in 2010 (except for NN, only 6 individuals in this population) for MLST (398 individuals in total). The sequences of the MLST genes (gatB, coxA, icp10, ftsZ, and fbpA) were amplified with standard primers and protocols (Baldo et al. 200691). After verification by gel electrophoresis, the PCR templates were purified, then sequenced in both directions using the five primer pairs respectively on an Applied Biosystems 3130 Genetic Analyzer.

mtDNA sequence analyses. DNA sequences were assembled and aligned with CodonCode Aligner 3.6.1 (CodonCode, Dedham, MA, USA), and manually edited before creating consensus sequences. To eliminate missing data, all sequences of the COI and COII genes were truncated to the same length of 769 bp and 599 bp, respectively, and the truncated sequences of the two genes were then concatenated into a data matrix for subsequent analysis. In order to assess the genetic variation within populations, the number of haplotypes (Nn), nucleotide diversity (π), and haplotype diversity (Hd) were calculated with DNASP v5.10.01. Shapiro-Wilk tests implemented in R software indicated that distributions of Hd and π from each respective climatic zone (SS, NS, WT and MT) were normal. T-tests were thus used to evaluate the significance of differences in diversity estimates Hd and π of the female populations between the four climatic zones and the P values were corrected by the FDR correction method.

The degree of population differentiation was quantified using pairwise FST values in Arlequin 3.5. Genotypic differentiation between pairwise populations was detected by an exact test. In order to visualize these genetic relationships between populations, the pairwise FST matrix was then used to construct a principal coordinates analysis (PCoA) in GenAlEx 6.5. To detect an isolation-by-distance (IBD) effect, we compared the FST/G1 (1 – FST) matrix with a geographic distance matrix (mkm) using Mantel test implemented in BDMWS 3.16, with significance tests performed over 104 permutations. The geographic distances between sampling sites were computed based on the GPS coordinates. For the four sites that were sampled twice, we used only the first year of collection. A Reduced Major Axis (RMA) regression was used to evaluate the slope and intercept of the isolation-by-distance relationship. The genealogical relatedness between haplotypes was represented first through a median-joining network constructed with Network 4.6, and the phylogenetic relationships between haplotypes were also inferred by Bayesian inference (BI) using MrBayes 3.1.2. N. hagens was added as an outgroup (GenBank Accession No. JX880069). The best model of nucleotide substitution was determined based on the Bayesian Information Criterion (BIC) calculated by jModelTest 2.1.2. The Markov chain was run for 10 million iterations using four chains with a sampling period of 100 generations, and we discarded the initial 20% as burn-in. As we found that the 107 haplotypes were divided into two haplogroups, we calculated the mean P-distance between the two haplogroups using MEGA 5.0 software.

Demographic history was explored taking three different approaches. First, Tajima’s D0 and Fu’s Fs8 were used to test for neutrality. Second, the mismatch distribution of pairwise sequences was calculated to identify the demographic history, which was characterized by the mismatch distribution modal and Harpending’s raggedness index105. Both the neutrality test and mismatch distribution analysis were conducted in Arlequin 3.5 with 104 simulated samples. Finally, a Bayesian skyline plot (BSP) analysis was performed with BEAST v1.7.5. We adopted the piecewise-linear skyline model and 10 groups for Bayesian skyline coalescent tree priors. The MCCM chains were run with 107 iterations, sampling every 106 iterations and discarding the initial 20% as burn-in. The relaxed clock with uncorrelated lognormal distribution was used, which allowed rate variation between branches. Default values were used for all other parameters. The convergence of the chain was checked by the TRACER v1.5 program (http://tree.bio.ed.ac.uk/software/tracer) to ensure that effective sample sizes were above 200.

As we observed a non-random distribution of the mtDNA haplogroups, we tested for an association between haplogroup frequencies and climatic variables (PC1 and PC2) using logistic regressions. We first carried out two univariate logistic regressions to test for associations between the frequencies of the two mitochondrial haplogroups and the environmental variable PC1 using the SAM program. We considered a correlation as significant only when two logistic regression tests (LRTs, G and Wald tests) rejected the null hypothesis of no association between the genetic and the environmental variables (at the 0.1% level) after FDR correction. Partial Mantel tests were then performed in ZT software to evaluate the specific contribution of PC1 or PC2 to the distribution of mtDNA haplogroups. This complementary approach made it possible to evaluate the effect of PC1 on the distribution of mtDNA haplogroups while removing the variance explained by the IBD effect. We also performed partial Mantel tests between mtDNA haplogroup frequencies and PC1 or PC2 while controlling effects for geographic distances (Log Km). The Euclidean distance metrics of haplogroup I (HGI) frequencies and environmental variables (PC1 and PC2) were calculated by ‘ecodist’ implemented in R 3.0.1.

Microsatellite data analyses. The population genetic diversity indices, such as the number of alleles (Na), observed heterozygosity (Ho), and unbiased expected heterozygosity (He) were assessed using GenAlEx 6.5. Allelic richness (Ar) was estimated using FSTAT 2.9.3.2.104. A rarefaction index corrected for different sample sizes, and the NN population was excluded owing to its small population size. In addition, we used Wilcoxon signed-rank tests applied on the 13 single locus values of the He and A0 averaged over all populations of the four climatic zones to determine whether genetic diversity within populations differed significantly between climatic zones and the P values were corrected by the FDR correction method.

Pairwise FST values were calculated in Arlequin v3.5. Genotypic differentiation between pairwise populations was detected by an exact test. The microsatellite loci employed in this study had very high within-population heterozygosity, which tends to provide low estimates of FST. To obtain a more accurate estimate of population differentiation, we calculated FST/G1 (1 – FST)105. To calculate the FST/G1, we calculated i = 1, . . . , N, where FST/G1 values were subsequently used for principal coordinate and IBD analyses. We used the individual clustering approach implemented in Structure v2.3.4 software to infer the number of functional population units. Simulations were run using the admixture ancestry model and the option of correlated allele frequency between populations. As one of the requirements of this collection scheme involved the selection of distant and at least 1 km apart locations, we used the USEPOPINFO option to incorporate sampling locations as prior information (LOCPRIOR)109. We set the number of genetic clusters from K = 1–12, because InPD values always decreased when approaching K = 12 (see
60. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. A-Stat. Soc.* 57, 289–300 (1995).
61. Chapuis, M. P. & Estoup, A. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* 24, 621–631 (2007).
62. Zhou, W., Rousset, F. & O’Neill, S. Phylogeography and PCR-based classification of *Wolbachia* strains using wsp gene sequences. *Proc. R. Soc. B* 265, 509–515 (1998).
63. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–95 (1989).
64. Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–25 (1997).
65. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–569 (1992).
66. Hedrick, P. W. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53, 313–318 (1999).
67. Meirmans, P. G. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60, 2399–2402 (2006).
68. Hubisz, M. J., Falush, D., Stephens, M. & Pritchard, J. K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* 9, 1322–1332 (2009).

Acknowledgments

We thank Da-Song Chen, Zi-Wei Song and Wen-Chao Zhu of the Department of Entomology, Nanjing Agricultural University (NJAU) for help with the collection of SBPH. We are also grateful to Dr. Feng Zhang of the Department of Entomology, NJAU for their kind help with data analyses. We thank Peter Biggins (CIRAD Scientific and Technical Information Service) for careful English language editing. This work was supported by a grant-in-aid from the National Key Basic Research Program (973 Program, No. 2009CB119200) from the Ministry of Science and Technology of China, a grant-in-aid (No. 200903051) from the Science and Technology Research Program of the National Agricultural Public Welfare Fund from the Ministry of Agriculture of China and a grant-in-aid from the Specialized Research Fund for the Doctoral Program of Higher Education (SRDFP) from the Ministry of Education of China (Priority Development Area, 20110907130005).

Author contributions

Conceived and designed the experiments: J.T.S. and X.Y.H. Performed the experiments: J.T.S., M.M.W., X.Y.J., Y.K.Z., C.G. and X.M.Y. Analyzed the data: J.T.S., M.P.C., G.H. and X.F.X. Wrote the paper: J.T.S., M.P.C. and X.Y.H.

Additional information

Data Accessibility: GPS values of the sampling locations, microsatellite genotype data, data matrix of the concatenated partial COI and COII gene sequences from 1323 individuals, and the 107 haplotype sequences of the concatenated partial COI and COII gene sequences are all deposited at Dryad (DOI:10.5061/dryad.6692, http://datadryad.org/review?wID=36983&token=18a66b98-ad73-4d2c-99e0-850da33c67d7).

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sun, J.-T. et al. Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, *Laodelphax striatellus*. *Sci. Rep.* 5, 8045; DOI:10.1038/srep08045 (2015).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/