



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Chen, MZ;Cao, LJ;Li, BY;Chen, JC;Gong, YJ;Yang, Q;Schmidt, TL;Yue, L;Zhu, JY;Li, H;Chen, XX;Hoffmann, AA;Wei, SJ

Title:

Migration trajectories of the diamondback moth *Plutella xylostella* in China inferred from population genomic variation

Date:

2021-04-01

Citation:

Chen, M. Z., Cao, L. J., Li, B. Y., Chen, J. C., Gong, Y. J., Yang, Q., Schmidt, T. L., Yue, L., Zhu, J. Y., Li, H., Chen, X. X., Hoffmann, A. A. & Wei, S. J. (2021). Migration trajectories of the diamondback moth *Plutella xylostella* in China inferred from population genomic variation. *Pest Management Science*, 77 (4), pp.1683-1693. <https://doi.org/10.1002/ps.6188>.

Persistent Link:

<https://hdl.handle.net/11343/276662>

Schmidt Thomas (Orcid ID: 0000-0003-4695-075X)

Chen Xue-xin (Orcid ID: 0000-0002-9109-8853)

Wei Shu-Jun (Orcid ID: 0000-0001-7398-0968)

Migration trajectories of the diamondback moth *Plutella xylostella* in China inferred from population genomic variation

Ming-Zhu Chen^{1,2}, Li-Jun Cao¹, Bing-Yan Li^{1,2}, Jin-Cui Chen¹, Ya-Jun Gong¹, Qiong Yang³, Thomas L. Schmidt³, Lei Yue¹, Jia-Ying Zhu⁴, Hu Li², Xue-Xin Chen⁵, Ary Anthony Hoffmann³, Shu-Jun Wei^{1,*}

1. Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
2. Department of Entomology and MOA Key Lab of Pest Monitoring and Green Management, College of Plant Protection, China Agricultural University, Beijing 100193, China
3. School of BioSciences, Bio21 Institute, The University of Melbourne, Victoria 3010, Australia
4. Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, College of Forestry, Southwest Forestry University, Kunming 650224, China
5. State Key Laboratory of Rice Biology and Ministry of Agriculture Key Laboratory of Agricultural Entomology, Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China

***Corresponding author:** Shu-Jun Wei (ORCID: 0000-0001-7398-0968)

Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China; E-mail: shujun268@163.com; Tel: +86 010-51503439; Fax: +86 010-51503899

Running title: Migration trajectories of the diamondback moth

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/ps.6188](https://doi.org/10.1002/ps.6188)

This article is protected by copyright. All rights reserved.

Abstract

BACKGROUND:

The diamondback moth (DBM), *Plutella xylostella* (Lepidoptera: Plutellidae), is a notorious pest of cruciferous plants. In temperate areas, annual populations of DBM originate from adult migrants. However, the source populations and migration trajectories of immigrants remain unclear. Here, we investigated migration trajectories of DBM in China with genome-wide single nucleotide polymorphisms (SNPs) genotyped using double-digest RAD (ddRAD) sequencing. We first analyzed patterns of spatial and temporal genetic structure among southern source and northern recipient populations, then inferred migration trajectories into northern regions using discriminant analysis of principal components (DAPC), assignment tests and spatial kinship patterns.

RESULTS:

Temporal genetic differentiation among populations was low, indicating sources of recipient populations and migration trajectories are stable. Spatial genetic structure indicated three genetic clusters in the southern source populations. Assignment tests linked northern populations to the Sichuan cluster, and central-eastern populations to the South and Yunnan clusters, indicating that Sichuan populations are sources of northern immigrants and South and Yunnan populations are sources of central-eastern populations. First-order (full-sib) and second-order (half-sib) kin pairs were always found within populations, but about 35-40% of third-order (cousin) pairs were found in different populations. Closely related individuals in different populations were in about 35-40% of cases found at distances of 900 to 1500 km, while some were separated by over 2000 km.

CONCLUSION:

This study unravels seasonal migration patterns in the DBM. We demonstrate how careful sampling and population genomic analyses can be combined to help understand cryptic migration patterns in insects.

Keywords: Seasonal migration, population genomics, individual assignment, kinship analysis

1 Introduction

The diamondback moth (DBM), *Plutella xylostella* (Lepidoptera: Plutellidae), is a notorious pest of cruciferous vegetables and oilseed rape¹⁻³. The annual cost of control and yield losses due to DBM has been estimated at US\$4-5 billion worldwide⁴. In most regions with temperate climates, DBM cannot survive cold winters in the field⁵⁻⁷. Annual populations of DBM in these areas originate from seasonal adult migrations⁸⁻¹⁰, which makes DBM a challenging pest to control². Understanding the migration routes of DBM is essential to improve outbreak forecasts and management of this pest¹¹. For example, when source populations are known, pest monitoring and control can be conducted in the source areas to reduce moth numbers migrating to temperate areas.

Previous studies have investigated migration of DBM by using field surveys, radar monitoring, light-trapping and laboratory tests^{9, 12-14}, and genetic approaches¹⁵⁻¹⁸. Field studies mainly monitored the dynamic of immigrants^{9, 13, 19}, while laboratory studies investigated the flight ability of the DBM^{12, 20}. Population genetic studies showed that populations of DBM are structured among continents¹⁸ but panmictic within a continent^{15-17, 21}. These studies also suggested that migration of DBM occurred within a continent but rarely across continents^{15, 18}. However, many aspects of DBM migration remain unclear, particularly in tracing source locations of migratory populations within continents and inter-annual variation in migration trajectories^{2, 3, 22}.

Migration is a life history strategy that has evolved in a large number of species across a diverse array of taxonomic groups as an adaptation to a changing environment^{23, 24}. Many animals migrate annually from their breeding areas to new habitats, to escape inclement conditions, and to exploit new resources²⁵⁻²⁷. The largest group of migratory animals are insects, which include both pests and natural enemies^{22, 28}. Some insects undertake long-distance seasonal migration with repeated movement between breeding and nonbreeding areas, as also found in a range of animals including birds and whales²⁹. However, many other insect species migrate opportunistically or without an overall pattern of cyclicity in return movement³⁰. Migration pathways can also be modified by environmental contingencies and natural selection^{25, 31}. These features of insect migration make it one of

the most difficult phenomena to study^{25, 32}. Additionally, establishing insect migration routes can be challenging because insect migration often occurs nocturnally and at high altitudes, and because the small size of insects means that they cannot be easily tracked using standard observational approaches^{14, 33-36}.

Genetic approaches have been used to trace the historical and recent movements of species across geographical scales³⁷⁻⁴⁰ and thus provide a powerful and efficient method of studying migration trajectories⁴¹⁻⁴⁴. These approaches also have challenges. Migration pathways can often only be inferred when populations are genetically differentiated, and thus can be difficult when gene flow is high^{23, 41, 45}. Even if genetic differentiation exists across broad geographical regions, it may not occur within regions, preventing the identification of source and recipient populations at finer scales^{22, 41, 46}. Also, temporal changes in source population genetic structure and hybridization among immigrants after their arrival may obscure the relationship between source and recipient populations^{23, 44}. Genetic studies of migration can also be hampered when a limited number of genetic markers are used, and when samples are collected at random time points across space rather than being timed to coincide with the migration event. We can improve our understanding of the migration trajectories of small insects by using population genomic approaches that integrate spatially and temporally constrained sampling, by using large numbers of genome-wide genetic markers, and by applying methods for analyzing kinship and assigning individuals to related groups^{37, 47}.

In this study, we use population genomic approaches to uncover the migratory trajectory of DBM. To overcome issues associated with high mobility and weak population genetic differentiation of DBM within continents¹⁵⁻¹⁷, we (1) used an improved sampling strategy by capturing immigrants from recipient populations immediately after their arrival for two successive years; (2) employed high resolution, genome-wide single nucleotide polymorphism (SNP) markers; and (3) used recently developed analytical methods in population genomics to link potential source populations and immigrants of DBM. We aimed to trace the source populations of the DBM annually migrating to northern China and to test the power of combining strategic sampling and genomic approaches to advance our understanding of insect migration. Based on previous population genetic structure studies,

we hypothesized that DBMs in southern China are consistently genetically distinguishable, and that immigrants entering northern China originated either from one or several of these genetically differentiated sources in southern China.

2 Materials and methods

2.1 Specimen collection and DNA extraction

DBM were sampled from potential source population locations in the annual breeding area of southern China. DBM were collected from cabbage and oilseed rape fields, and all sampling was completed before the first observations of DBM in northern China between March and May^{13, 48}. In order to reduce the likelihood of sampling siblings within populations, third- and fourth-instar larvae of DBM were collected from about 20 sites at each sampling location, each at least 10 m apart. Putative immigrant male adults were collected in northern China by sex pheromone trapping before the presence of first-generation larvae. Trapping of male DBM was conducted in unplanted fields with no greenhouses within 500 m, to reduce the likelihood of trapping individuals overwintering in protected conditions. The distance between traps was at least 50 m. The development of one generation of DBM takes about 30 days in early spring². This strategy therefore restricted sampling of genetically related individuals to within three generations between source and recipient populations, and reduced the influence of genomic admixture between immigrants from different sources. This sampling was conducted in 2017 and again in 2018, to examine annual variation in migratory trajectories and temporal variation in population genetic structure. In total, samples were collected from 16 locations in 2017 and 17 locations in 2018, and in 2018 four locations were sampled across multiple months (Fig. 1, Table 1). Twenty individuals from each population (specimens collected at different times from the same location were considered as different populations) were used for genotyping. Genomic DNA for library preparation was extracted from individual specimens using DNeasy Blood and Tissue Kit (Qiagen, Germany).

2.2 SNP genotyping

The ddRAD libraries were prepared following a published protocol⁴⁹ for identifying SNPs. Briefly, 120 ng of extracted genomic DNA from each sample was digested by the restriction

enzymes *NlaIII* and *Acil* (New England Biolabs, USA) ⁵⁰. The 50 μ L digestion reaction was run for 3 hours at 37 °C, followed by DNA cleaning using 1.5 \times volume of AMPure XP beads (Beckman Coulter, USA) instead of a heat kill step. Next, we ligated each sample to adapters barcoded with a combinatorial index at 16 °C overnight in a 40 μ L ligation reaction, labeling each population with a 6-bp index and each individual with a unique 9-bp barcode. After ligation, we pooled uniquely barcoded samples into multiplexed libraries. Fragments between 380-540 bp were selected using BluePippin and a 2% gel cassette (Sage Sciences, USA). Finally, the pooled libraries were enriched with 12 amplification cycles on a Mastercycler Nexus Thermal Cycler (Eppendorf, Germany). PCR products were cleaned with 0.8 \times volume of beads. We used Qubit 3.0 (Life Invitrogen, USA) and Agilent 2100 Bioanalyzer (Agilent Technology, USA) to check the concentration and size distribution of enriched libraries, respectively. Pooled libraries were sequenced on an Illumina HiSeq 2500 platform to obtain 150-bp paired-end reads, at BerryGenomics Company (Beijing, China).

The Stacks v2.3 pipeline ⁵¹ was used to call SNPs, linking to the DBM genome (GenBank assembly accession: GCA_000330985.1) as reference ⁵². FastQC v 0.11.5 was employed to assess read quality and check for adapter contamination ⁵³. Sequence data was demultiplexed and trimmed using `process_radtags` in Stacks v2.3 ^{51, 54}. Low quality reads with a Phred score below 20 were removed as well as any reads with an uncalled base. Reads were trimmed to 140 bp in length. The remaining paired-end reads were aligned to the DBM genome ⁵² using Bowtie v2.3.5 ⁵⁵.

Output reads for all individuals were imported into Stacks pipeline `ref_map.pl` to call SNPs, requiring a minimum of three identical reads to create a stack. SNPs were called using a maximum likelihood statistical model. Finally, we obtained a catalog with all possible loci and alleles. The exported loci were present in all populations, and in at least 75% of individuals per population. The exported SNPs for populations that were collected in both years were further filtered using the R package `vcfR` ⁵⁶ and VCFtools v0.1.16 ⁵⁷ with the following criteria: SNPs with sequencing depth ≤ 3 and in the highest 0.1% depth were removed, as were SNPs with missingness in all samples ≥ 0.05 and those with minimum minor allele count ≤ 20 . An additional data matrix was generated by retaining only SNPs separated by at least 500 bp, to reduce linkage among SNPs.

2.3 Genetic diversity, population structure and assignment tests

Global population differentiation was estimated using Weir and Cockerham's F_{ST} with 99% confidence intervals (1000 bootstraps) in *diveRsity* version 1.9.90. Pairwise F_{ST} for all population pairs was estimated using GenePop version 4.7.2⁵⁸. Discriminant analysis of principal components (DAPC) was performed in the R package *adegenet* v2.1.1⁵⁹, with the optimal number of clusters determined by the Akaike information criterion (AIC).

Assignment tests were performed in *assignPOP* v1.1.7⁶⁰. Source groups of ST (south) and SW (southwest, this group was divided into YN and SC groups in 2018) (see Table 1 and Fig. 1 for locations) were trained using the support vector machine algorithm to build predictive models. For training, we used either 25, 28, or 32 random individuals (2017 samples) or 13, 15 or 17 random individuals (2018 samples) from each group, and loci with the highest 60%, 80% or 100% F_{ST} values. Monte-Carlo cross-validation was performed by resampling each training set combination 1000 times. The ratio of assignment probability between the most-likely and second most-likely assigned groups was calculated for each individual⁶¹. When an individual showed an assignment ratio smaller than 2 in more than 30% of the resampling analysis, it was considered unstable and removed in subsequent training. This allowed us to remove individuals from source populations that are not similar enough to other individuals in that source population, thus leaving a set of source populations each comprised of individuals distinctive from those in other populations. Immigrants from the CE (central) and NT (north) regions (see Table 1 and Fig. 1 for locations) were assigned to the trained groups using the support vector machine algorithm.

2.4 Kinship analysis

As a complement to assignment tests (but focusing on the individual level rather than the population level), we investigated spatial patterns of kinship within and between populations. Related individuals were identified following the method of Jasper, Schmidt, Ahmad, Sinkins and Hoffmann⁴⁷. First, Loiselle's K was calculated for all individual pairs using SPAGeDi⁶². Kinship coefficients represent the probability that any allele scored in both individuals is identical by descent, with theoretical mean K values for each kinship category as follows: full-siblings = 0.25, half-siblings = 0.125, full-cousins = 0.0625, half-

cousins = 0.0313, second-cousins = 0.0156 and unrelated = 0. To allocate pairs of individuals to relatedness categories across three orders of kinship, maximum-likelihood estimation in the program ML-Relate⁶³ was used to identify first-order (full-sibling) and second-order (half-sibling) pairs. The K scores of pairs within the full-sibling and half-sibling data sets were used to calculate standard deviations for these categories. Using the theoretical means and standard deviations of K , we randomly sampled 100,000 simulated K scores from each kinship category. In the initial pool of 40755 pairings (2017) and 89676 pairings (2018), ML-Relate identified 33 (2017) and 36 (2018) full-sibling and half-sibling pairs. Assuming that the data contained twice as many first cousin (full and half) pairings as sibling (full and half) pairings, and twice as many second cousin pairings as first cousin pairings, final sampling distributions were developed as follows: 100,000 unrelated, 320 second-cousins, 80 full-cousins, 80 half-cousins, 40 half-siblings, 40 full-siblings (2017) and 100,000 unrelated, 160 second-cousins, 40 full-cousins, 40 half-cousins, 20 half-siblings, 20 full-siblings (2018). To analyze how closely this distribution approximates the field data, we randomly sampled 40,000 (2017) and 80,000 (2018) simulated K scores from the above sampling distribution and plotted a histogram of this combined distribution and a histogram of the unrelated distribution against a histogram of 40,000 (2017) and 80,000 (2018) K scores from the empirical data. As the combined distribution matched the empirical distribution much more closely than the unrelated distribution, we adopted it for kinship inference and used it to determine appropriate kinship coefficient cutoffs for third-order (cousin) relationships, following Jasper, Schmidt, Ahmad, Sinkins and Hoffmann⁴⁷.

3 Results

3.1 Genotyping potential source populations and immigrants

When populations of both years were analyzed together, we retained 12293 SNPs after filtering. When collection years were analyzed separately, we retained 5628 SNPs for the 286 individuals collected in 2017, and 8750 SNPs for the 424 individuals collected in 2018. We filtered these SNPs to retain only SNPs separated by at least 500 bp to reduce the effects of linkage among loci on genetic diversity and genetic structure estimates. This left

2100 SNPs when both collection years were combined, 1580 SNPs from 2017 alone, and 2339 SNPs from 2018 alone (Table S1).

Estimates of genetic diversity parameters were consistent across populations. Allelic richness (A_R) ranged from 1.32 to 1.41, the percentage of total alleles (%Poly) ranged from 70.31% to 82.60% and the average nucleotide diversity (P_i) estimated for each population ranged from 0.092 to 0.103. All populations deviated from Hardy-Weinberg equilibrium (HWE) with an apparent inbreeding coefficient (F_{IS}) ranging from 0.142 to 0.221 (Table 1), as is typical for insects analyzed with SNPs^{38,47}.

3.2 Significant differentiation across spatial populations but not among temporal populations

Global differentiation of DBM was low, with Weir and Cockerham's F_{ST} having a value of 0.0104 (95% confidence intervals, CI: 0.0065 - 0.0142) for populations from both years, 0.0106 (95% CI: 0.0046 - 0.0179) for populations from 2017, and 0.0095 (95% CI: 0.0046 - 0.0148) for populations from 2018. The pairwise F_{ST} among populations from the two years ranged from -0.0033 to 0.0347 with quartiles of 0.0049 (25%) and 0.0150 (75%). Populations from southern China (HNHK, GDGZ and GDSZ) showed relatively higher-level genetic differentiation from other populations as reflected by higher F_{ST} values (Fig. S1).

Samples collected at different times from the same location in 2018 showed the lowest level of differentiation, including those collected from southern areas (F_{ST} among GDGA, GDGB and GDGC: -0.0033 - 0.0006; among JXNA, JXNB, JXNC: -0.0021 - 0.002) and northern areas (among BJYA, BJYB, BJYC: -0.0001 - 0.0008; between HBZA and HBZB: 0.0022) in 2018. Populations collected in different years from the same area also had relatively low levels of differentiation, including populations collected from southern areas (F_{ST} between HNHK in 2017 and 2018: 0.0013; between GDGZ in 2017 and GDGA, GDGB, GDGC in 2018: 0.0015 - 0.0047; between SCCD in 2017 and 2018: 0.0069; between SCGY in 2017 and 2018: 0.0073; between JXNC in 2017 and JXNA, JXNB, JXNC in 2018: 0.0019 - 0.0050) and northern areas (between BJYQ in 2017 and BJYA, BJYB, BJYC in 2018: -0.0002 - 0.0013; between HBZJ in 2017 and HBZA, HBZB in 2018: -0.0007 - 0.0016) (Fig. S1). These values show that there is

relatively little genetic change over time at locations in both southern and northern China (Fig. S1).

Mantel tests revealed significant correlation between genetic and geographical distances among populations in 2018 ($r = 0.335$, $p = 0.001$) and among populations across both years ($r = 0.294$, $p = 0.001$), but not among populations collected in 2017 ($r = 0.163$, $p = 0.115$). When putative immigrant populations from northern area were excluded from the 2017 analysis (see below), a significant association between genetic and geographic distance was also found for the 2017 samples ($r = 0.299$, $p = 0.013$) (Fig. S2).

3.3 Genetic structure analysis identified major clusters among potential source populations

To examine seasonal migration, we first inferred the population genetic structure of the DBM using discriminant analysis of principal components (DAPC)⁵⁹, which does not rely on the assumption of HWE within populations. When populations from the two years were analyzed together, we identified two large genetic clusters: southern populations were assigned to one cluster regardless of their collection year, while the other populations were assigned to the second cluster (Fig. S3). These results suggest population genetic structure of potential source and recipient populations is relatively stable over time.

For populations collected in 2017, three clusters were identified: the large cluster 1 was composed of individuals mainly from populations collected from southwestern, central-eastern (CE) and northern China; the large cluster 2 was composed of individuals mainly collected from two southern populations and two populations from CE (JXNC and HBWH); and the small cluster 3 was composed of some individuals from a southwestern population (Fig. 2a, 2b). A similar population genetic structure was found in 2018 with four clusters identified: the southwestern cluster in 2017 was further divided into cluster 1, which was composed of individuals mainly from Sichuan and northern populations, and cluster 2, which was composed of individuals mainly from Yunnan and CE; cluster 3 was composed of individuals mainly from the southwestern area as in 2017, while the small cluster 4 was composed of some individuals from a population of the Yunnan group (Fig. 2c, 2d). The results indicate that individuals from northern populations overlapped with putative source populations. This involved overlap of northern populations mainly with populations from

southwestern China in 2017, and with Sichuan in 2018. Central and eastern populations (CE) mainly overlapped with those from southwestern China in 2017, and from Yunnan in 2018 (Fig. 1). These shifting patterns also emphasize that the northern populations are indeed composed of migrants.

3.4 Population connectivity revealed from individual assignment tests

Having established the genetic structure of the potential source populations, we performed assignment tests on individuals from populations of the CE and NT groups, using a machine learning method implemented in *assignPOP*⁶⁰. Source groups of ST and SW (divided into YN and SC in 2018) were trained using the support vector machine algorithm to build predictive models. After cross-validation to remove the individuals that were difficult to assign, 35 and 92 individuals remained in the ST and SW groups respectively in 2017, and 37, 22 and 19 individuals remained in the ST, YN and SC groups respectively in 2018. On average, 78.54% of CE and 95.62% of NT were assigned to SW in 2017; nearly half of the CE group individuals were assigned to YN (49.79%), while 92.81% of the NT individuals were assigned to YN (44.71%) and SC (48.10%) in 2018. When individuals with a proportion of the genetic constitution of one source > 75% were considered as an effective assignment, 70.21% of individuals of CE were assigned to SW, while 93.55% of individuals of NT were assigned to SW in 2017. In 2018, 67.86% of individuals of CE and 95.80% of individuals of NT were assigned to SW, but it was not possible to assign these further to YN or SC (Table S2, Fig. 3).

3.5 Trajectories and distance of migration revealed from spatial kinship distributions

Using ML-Relate⁶³, 12 full sibling pairs and 21 half-sibling pairs were identified in 2017, while 17 full sibling pairs and 19 half-sibling pairs were identified in 2018. All of these were found within populations. Full siblings were from the same field. After simulations following the method of Jasper, Schmidt, Ahmad, Sinkins and Hoffmann⁴⁷, we defined Loiselle's $K = 0.071$ (2017) and $K = 0.063$ (2018) as cutoffs for describing cousin relationships (Fig. S4). Kinship analysis using SPAGeDi⁶² showed that 0.18% and 0.11% pairs of individuals were related at this level when all pairs of individuals from 2017 and 2018 were considered respectively. In 2017, 34 pairs of putative cousins were identified, while in 2018, 54 pairs were identified. Of these, 64.71% (2017) and 59.26% (2018) represented cousins from the

same population collected in the same location. Other related individuals tended to occur in populations from nearby groups (Fig. 4a, 4e). We also examined putatively related pairs of individuals that had the top 1% or 10% of K values (Table S3). When the 10% cutoff was used, 44.77% of the related individual pairs in 2017 occurred between the ST and CE groups, between the SW and CE groups, and between the SW and NT groups. In 2018, 47.25% of the related individual pairs occurred between CE and the NT, ST, YN groups, and between NT and the YN, SC groups. Considering these in terms of geographical distance, 42.39% (2017) and 33.72% (2018) of related individual pairs were found in populations separated by distances of 900- 1500 Km; related individuals more than 2400 Km apart comprised 0.29% and 0.91% of the related pairs in 2017 and 2018, respectively (Fig. 4b-d, 4f-4h).

4 Discussion

4.1 *The southwestern populations as the main sources for immigrants of DBM in northern China*

Long-distance migration in DBM had already been documented in the 1860s in Europe⁶⁴ and in the 1930s in the U.S.⁶⁵, and is a common phenomenon worldwide, including in Asia and Australia^{13, 15, 66}. Early studies reported migration of DBM based on the sudden appearance of large numbers of moths in the spring and the capture of specimens over the sea far from land⁶⁷. Recent studies monitored DBM migration using radar in introduced areas and characterized aspects of migratory behavior^{9, 13}. However, linking immigrants to specific source populations by back-tracking methods was not possible, due to widespread breeding habitats in subtropical and tropical regions².

We used a population genomic approach incorporating assignments tests and kinship inference to unravel migration trajectories of DBM across China. This region covers both temperate areas (northern China) where DBM cannot survive the cold winter in the field, and tropical and subtropical areas (southern China) where DBM can breed year-round³. Previous studies showed a lack of genetic differentiation among populations across China but apparent population genetic differentiation among continents^{15, 17, 18, 66}. Observations on DBM populations suggest DBM most likely migrates within rather than between continents⁶⁷. We also found that related individuals linked to the southern source

populations tended to be separated by geographical distances of 900 to 1500 km rather than larger distances. However occasional movement across larger distances was detected.

The genetic structure of DBM from the same location and area in southern China was relatively constant, as reflected by low F_{ST} values within populations over time. This suggests that there was no large-scale movement of immigrants into southern Chinese populations from other regions in the two years of this study. We did find four individuals collected in 2017 and three individuals collected in 2018, all from southwestern populations, which were assigned to unique clusters, which may reflect introductions of DBM from other isolated populations. Long distance movement of small invertebrates can occur through passive transport, which has been well documented for mosquitos transferred through aircraft and thrips transferred through movement of plants⁶⁸. Nevertheless, these movements appear relatively rare in DBM, unlike for some other insects in southern China such as planthoppers breeding on rice, for which there are substantial northward migrations from other southeastern countries⁶⁹.

Among the source populations, we found two genetic clusters in 2017 and three clusters in 2018. In 2018, the southwestern cluster from 2017 was further divided into Sichuan and Yunnan clusters. This different pattern between the years is probably just a reflection of sampling density. When we noticed that southwestern populations were linked to northern populations in 2017, we increased sampling density around Yunnan in 2018 (Fig. 1), increasing our ability to separate these clusters.

The predicted worldwide distribution of DBM based on a validated bioclimatic model shows that DBM can persist year-round in all collected regions of our study except for northern China². Population genetic structure, assignment and kinship analysis indicated that immigrants of the DBM in northern China mainly came from southwestern areas, specifically around the Sichuan area. We also found that populations collected from central and eastern China were a mixture from southern and southwestern (specifically around Yunnan province) populations, with the latter as a primary source. Our results suggest that populations from southern China are relatively isolated from other populations and provide immigrants mainly to central and eastern populations but less frequently to northern China. Populations of DBM in south China have the highest levels of pesticide resistance^{3,70}, but a

high frequency of pesticide application still reduces the density of DBM in this area where cabbage is often planted³. Lower levels of pesticide resistance of DBM from northern China compared to south China³ also support the notion that southern populations are not the main source of migrants moving into northern regions.

Moths experience a greater degree of wind drift than migrating songbirds, allowing for long distance movement but at the cost of a comparatively lower spatiotemporal migratory precision⁷¹. The abundance of migratory DBM may be associated with winds from storm fronts¹⁵. The East Asian subtropical monsoon could facilitate migration of this insect to northern China. The Sichuan area, which provided the main sources of immigrants to northern China, has large areas of cultivated oilseed, which may allow high numbers of DBM to accumulate before migration. The harvest time of oilseed in Sichuan area is coincident with the first peak of DBM in northern China from April to May⁴⁸. Source populations of DBM also need suitable climatic conditions as well as host plants^{2,9}, and Shirai²⁰ found that adult DBM developing in cooler seasons are well adapted to migration. This could also favor Sichuan as a source area because southern areas have higher temperatures than Sichuan prior to the period of northward migration. Nevertheless, studies conducted over a larger number of years could help to monitor the annual variation in source populations and migration trajectories of DBM as climates and cultivation systems change.

Migration distances vary in different groups and species. The migration distance of moths can exceed over 3000 km⁷², while birds can migrate farther than 10000 km⁷³. DBM migration distance estimates have varied considerably across studies, from less than 110 m⁷⁴, to around 300 km⁷⁵, and over 1000 km⁷⁶. Migration of DBM from southern Europe to the UK⁹ points to distances of 1500 km which is comparable to the migration distances of 900 km to 1500 km observed here. The distances we have uncovered may reflect stepping-stone movement across multiple generations, given that the sampling used in our study could have encompassed three generations. Variation in flight distance may result from many variables like nutrition, weather conditions and pathogen load⁷³.

4.2 Population genomic approaches incorporated with strategic sampling provide efficient methods to study animal migration

Animal migration usually occurs over broad scales of time and space relative to the size and development of the migrant, which makes it challenging to investigate²⁵. Numerous methods from multiple disciplines have been employed to investigate migration²⁵. By far the most widely used approach is to track migrant animals with passive extrinsic markers, but this is mainly suitable for large animals rather than for small insects⁷⁷. Radar is one of the promising tools for monitoring the migration behaviors of small insects and large animals^{14, 78, 79}. To be useful, this method needs large-scale and long-term radar networks and improvements in target recognition, especially for small insects^{80, 81}.

Genetic approaches have been proposed as efficient methods to trace migrant sources; however, applying these in the wild has proved challenging^{41, 66}. Previous population genetic studies investigating migratory trajectories of DBM have used marker sets such as mitochondrial genes and microsatellites, but these have not shown much differentiation among populations¹⁵. A ddRADseq approach has allowed us to genotype large numbers of individuals at thousands of genome-wide SNPs, and this approach can help to distinguish weakly differentiated populations of DBM^{15, 82}. SNP based approaches are starting to become widely used in ecological and evolutionary studies and provide novel insights into population processes^{83, 84}. In our study, we found that the vastly increased number of genetic markers helped us to identify a weak but clear pattern of genetic differentiation among some populations of DBM. We also found a high level of genetic diversity in populations, as expected for a migratory species⁸⁵.

Sampling strategy can influence the outcome of population genetic analyses and consequent ecological inferences⁸⁶. Estimates of population divergence will be influenced by the timing of sampling in relation to migration events⁴⁵. Unlike earlier work on DBM, we used a specific strategy of sample collection to identify migratory trajectories of DBM, collecting DBM from the likely source and target populations within a short time. Putative immigrants were collected immediately after their arrival. Our sampling strategy probably helped to avoid the collection of any hybrids from different sources, and maximized the likelihood of collecting genetically related individuals migrating into an area. Development of this strategy benefitted from the well-studied biology of DBM and prior knowledge of its seasonal migration pattern¹⁻³. These methods of sampling, genotyping and population

genomic analysis could be readily used to uncover migration trajectories and discover migration behaviors of the DBM in other regions that suffer outbreaks related to immigration, and other small insects more broadly.

5 Conclusion

Despite considerable advances in understanding the biology, ecology and genetics of DBM, this moth remains the most destructive pest of *Brassica* vegetable crops globally¹⁻³. Applications of insecticides represent the most common strategy for controlling DBM⁷, but DBM shows resistance to all pesticides used to control it, including *Bacillus thuringiensis* toxins^{87,88}. A lack of knowledge of the source of migrants has hindered the development of forecasting systems for DBM management³. Our results point to source areas for monitoring DBM to improve the forecasting in receiving areas and for targeted control efforts. Management of DBM in Sichuan areas could potentially reduce the number of migrants entering northern China. We found that some individuals from southern areas, where level of pesticide resistance are usually high³, can migrate to northern China. Thus, monitoring the proportion of immigrants from southern China and developing genetic markers linked to resistance in southern populations may help to inform risk of pesticide resistance in northern China.

Acknowledgments

We are grateful to Moshe Jasper from The University of Melbourne for help in data analysis, and to Ming-Liang Li, De-Qiang Pu, Zhen-Qiang Qin, Qing-Sheng Lin, and Wei Ren for help on sample collection. This work was supported by the National Key Research and Development Program of China (2018YFD0201207), International Cooperation Foundation of Beijing Academy of Agriculture and Forestry Sciences (GJHZ2017-5), Beijing Key Laboratory of Environmentally Friendly Pest Management on Northern Fruits (BZ0432) and BAAFS-UOM Joint Laboratory on Pest Control Research.

Conflicts of interest

The authors declare no competing interests.

Data availability statement

Demultiplexed ddRAD sequencing reads will be available through the National Center for Biotechnology Information's Sequence Read Archive and all metadata and scripts associated with analyses in this study will be deposited to Dryad.

References

1. Talekar NS and Shelton AM, Biology, ecology, and management of the diamondback moth. *Annu Rev Entomol* **38**(275-301 (1993).
2. Furlong MJ, Wright DJ and Dosdall LM, Diamondback moth ecology and management: problems, progress and prospects. *Annu Rev Entomol* **58**(517-541 (2013).
3. Li Z, Feng X, Liu SS, You M and Furlong MJ, Biology, ecology, and management of the diamondback moth in China. *Annu Rev Entomol* **61**(277-296 (2016).
4. Zalucki MP, Shabbir A, Silva R, Adamson D, Liu SS and Furlong MJ, Estimating the economic cost of one of the world's major insect pests, *Plutella xylostella* (Lepidoptera: Plutellidae): just how long is a piece of string? *J Econ Entomol* **105**(4): 1115-1129 (2012).
5. Maeda M and Takano T, Seasonal prevalence and the possibility of overwintering of diamond-back moth *Plutella xylostella* (L.), in Miyagi prefecture. *Bulletin of the Miyagi Prefectural Horticultural Experiment Station* **5**(1-20 (1984).
6. Zalucki M and Furlong M. Predicting outbreaks of a migratory pest: an analysis of DBM distribution and abundance revisited. In *International Workshop on Management of the Diamondback Moth and Other Crucifer Insect Pests (6th, 2011) AVRDC: The World Vegetable Centre*, pp. 8-14 (2011).
7. Grzywacz D, Rossbach A, Rauf A, Russell DA, Srinivasan R and Shelton AM, Current control methods for diamondback moth and other brassica insect pests and the prospects for improved management with lepidopteran-resistant Bt vegetable brassicas in Asia and Africa. *Crop Prot* **29**(1): 68-79 (2010).
8. Hopkinson R and Soroka J, Air trajectory model applied to an in-depth diagnosis of potential diamondback moth infestations on the Canadian Prairies. *Agric For Meteorol* **150**(1): 1-11 (2010).
9. Chapman JW, Reynolds DR, Smith AD, Riley JR, Pedgley DE and Woiwod IP, High-altitude migration of the diamondback moth *Plutella xylostella* to the UK: a study using radar, aerial netting, and ground trapping. *Ecol Entomol* **27**(6): 641-650 (2002).
10. Honda K-i. Hibernation and migration of diamondback moth in northern Japan. In *The Management of Diamondback Moth and Other Crucifer Pests Proceedings of the Second International Workshop on Management of the Diamondback Moth and Other Crucifer Pests AVRDC-The World Vegetable Centre Tainan, Taiwan*, pp. 45-50 (1992).
11. Li Z, Zalucki MP, Bao H, Chen H, Hu Z, Zhang D, Lin Q, Yin F, Wang M and Feng X, Population dynamics and "outbreaks" of diamondback moth (Lepidoptera:

- Plutellidae) in Guangdong province, China: climate or failure of management? *J Econ Entomol* **105**(3): 739-752 (2012).
12. Shirai Y, Factors influencing flight ability of male adults of the diamondback moth, *Plutella xylostella*, with special reference to temperature conditions during the larval stage. *Appl Entomol Zoolog* **28**(3): 291-301 (1993).
 13. Fu X, Xing Z, Liu Z, Ali A and Wu K, Migration of diamondback moth, *Plutella xylostella*, across the Bohai Sea in northern China. *Crop Prot* **64**(143-149 (2014).
 14. Chapman JW, Drake VA and Reynolds DR, Recent insights from radar studies of insect flight. *Annu Rev Entomol* **56**(337-356 (2011).
 15. Endersby NM, McKechnie SW, Ridland PM and Weeks AR, Microsatellites reveal a lack of structure in Australian populations of the diamondback moth, *Plutella xylostella* (L.). *Mol Ecol* **15**(1): 107-118 (2006).
 16. Wei S, Shi B, Gong Y, Jin G, Chen X and Meng X, Genetic Structure and Demographic History Reveal Migration of the Diamondback Moth *Plutella xylostella* (Lepidoptera: Plutellidae) from the Southern to Northern Regions of China. *Plos One* **8**(4): e59654 (2013).
 17. Yang J, Tian L, Xu B, Xie W, Wang S, Zhang Y, Wang X and Wu Q, Insight into the migration routes of *Plutella xylostella* in China using mtCOI and ISSR markers. *PLoS One* **10**(6): e0130905 (2015).
 18. Roux O, Gevrey M, Arvanitakis L, Gers C, Bordat D and Legal L, ISSR-PCR: Tool for discrimination and genetic structure analysis of *Plutella xylostella* populations native to different geographical areas. *Mol Phylogenet Evol* **43**(1): 240-250 (2007).
 19. Campos WG, Schoereder JH and DeSouza OF, Seasonality in neotropical populations of *Plutella xylostella* (Lepidoptera): resource availability and migration. *Popul Ecol* **48**(2): 151-158 (2006).
 20. Shirai Y, Longevity, flight ability and reproductive performance of the diamondback moth, *Plutella xylostella* (L) (Lepidoptera: Yponomeutidae), related to adult body size. *Res Popul Ecol* **37**(2): 269-277 (1995).
 21. You M, Ke F, You S, Wu Z, Liu Q, He W, Baxter SW, Yuchi Z, Vasseur L, Gurr GM, Ward CM, Cerda H, Yang G, Peng L, Jin Y, Xie M, Cai L, Douglas CJ, Isman MB, Goettel MS, Song Q, Fan Q, Wang-Pruski G, Lees DC, Yue Z, Bai J, Liu T, Lin L, Zheng Y, Zeng Z, Lin S, Wang Y, Zhao Q, Xia X, Chen W, Chen L, Zou M, Liao J, Gao Q, Fang X, Yin Y, Yang H, Wang J, Han L, Lin Y, Lu Y and Zhuang M, Variation among 532 genomes unveils the origin and evolutionary history of a global insect herbivore. *Nature Communications* **11**(1): 2321 (2020).
 22. Jones CM, Parry H, Tay WT, Reynolds DR and Chapman JW, Movement ecology of pest *Helicoverpa*: implications for ongoing spread. *Annu Rev Entomol* **64**(277-295 (2019).
 23. Chapman JW, Reynolds DR and Wilson K, Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecol Lett* **18**(3): 287-302 (2015).
 24. Holland RA, Wikelski M and Wilcove DS, How and why do insects migrate? *Science* **313**(5788): 794-796 (2006).

25. Dingle H. *Migration: the biology of life on the move, second edition*. Oxford University, Oxford, United Kingdom, (2014).
26. Brower LP, Understanding and misunderstanding the migration of the monarch butterfly (Nymphalidae) in North America. *J Lepidopt Soc* **49**(304-385 (1995).
27. Zhan S, Merlin C, Boore JL and Reppert SM, The monarch butterfly genome yields insights into long-distance migration. *Cell* **147**(5): 1171-1185 (2011).
28. Hu G, Lim KS, Horvitz N, Clark SJ, Reynolds DR, Sapir N and Chapman JW, Mass seasonal bioflows of high-flying insect migrants. *Science* **354**(6319): 1584-1587 (2016).
29. Lyons JJ, Pierce AA, Barribeau SM, Sternberg ED, Mongue AJ and De Roode JC, Lack of genetic differentiation between monarch butterflies with divergent migration destinations. *Mol Ecol* **21**(14): 3433-3444 (2012).
30. Dingle H, Migration strategies of insects. *Science* **175**(4028): 1327 (1972).
31. Åkesson S, Flying with the winds: differential migration strategies in relation to winds in moth and songbirds. *J Anim Ecol* **85**(1): 1-4 (2016).
32. Cote J, Bocedi G, Debeffe L, Chudzińska ME, Weigang HC, Dytham C, Gonzalez G, Matthysen E, Travis J and Baguette M, Behavioural synchronization of large-scale animal movements-disperse alone, but migrate together? *Biol Rev* **92**(3): 1275-1296 (2017).
33. Gatehouse A, Behavior and ecological genetics of wind-borne migration by insects. *Annu Rev Entomol* **42**(1): 475-502 (1997).
34. Kays R, Crofoot MC, Jetz W and Wikelski M, Terrestrial animal tracking as an eye on life and planet. *Science* **348**(6240): aaa2478 (2015).
35. Bridge ES, Thorup K, Bowlin MS, Chilson PB, Diehl RH, Fléron RW, Hartl P, Kays R, Kelly JF and Robinson WD, Technology on the move: recent and forthcoming innovations for tracking migratory birds. *Bioscience* **61**(9): 689-698 (2011).
36. Nilsson C, Dokter AM, Verlinden L, Shamoun-Baranes J, Schmid B, Desmet P, Bauer S, Chapman J, Alves JA and Stepanian PM, Revealing patterns of nocturnal migration using the European weather radar network. *Ecography* **42**(5): 876-886 (2019).
37. Cayuela H, Rougemont Q, Prunier JG, Moore JS, Clobert J, Besnard A and Bernatchez L, Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Mol Ecol* **27**(20): 3976-4010 (2018).
38. Schmidt TL, van Rooyen AR, Chung J, Endersby-Harshman NM, Griffin PC, Sly A, Hoffmann AA and Weeks AR, Tracking genetic invasions: Genome-wide single nucleotide polymorphisms reveal the source of pyrethroid-resistant *Aedes aegypti* (yellow fever mosquito) incursions at international ports. *Evol Appl*: **12**(6): 1136-1146 (2019).
39. Ryan SF, Lombaert E, Espeset A, Vila R, Talavera G, Dincă V, Doellman MM, Renshaw MA, Eng MW and Hornett EA, Global invasion history of the agricultural pest butterfly *Pieris rapae* revealed with genomics and citizen science. *Proc Natl Acad Sci U S A* **116**(40): 20015-20024 (2019).

40. Perry KD, Keller MA and Baxter SW, Genome-wide analysis of diamondback moth, *Plutella xylostella* L., from Brassica crops and wild host plants reveals no genetic structure in Australia. *Scientific Reports* **10**(1): 12047 (2020).
41. Webster MS, Marra PP, Haig SM, Bensch S and Holmes RT, Links between worlds: unraveling migratory connectivity. *Trends Ecol Evol* **17**(2): 76-83 (2002).
42. Wink M, Use of DNA markers to study bird migration. *J Ornithol* **147**(2): 234-244 (2006).
43. Franchini P, Irisarri I, Fudickar A, Schmidt A, Meyer A, Wikelski M and Partecke J, Animal tracking meets migration genomics: transcriptomic analysis of a partially migratory bird species. *Mol Ecol* **26**(12): 3204-3216 (2017).
44. Larroque J, Legault S, Johns R, Lumley L, Cusson M, Renaut S, Levesque RC and James PMA, Temporal variation in spatial genetic structure during population outbreaks: Distinguishing among different potential drivers of spatial synchrony. *Evol Appl* **12**(10): 1931-1945 (2019).
45. Turbek SP, Scordato ESC and Safran RJ, The role of seasonal migration in population divergence and reproductive isolation. *Trends Ecol Evol* **33**(3): 164-175 (2018).
46. Anderson C, Tay WT, McGaughran A, Gordon K and Walsh TK, Population structure and gene flow in the global pest, *Helicoverpa armigera*. *Mol Ecol* **25**(21): 5296-5311 (2016).
47. Jasper M, Schmidt TL, Ahmad NW, Sinkins SP and Hoffmann AA, A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. *Mol Ecol Resour* **19**(5): 1254-1264 (2019).
48. Zhu LH, Li ZY, Zhang SF, Xu BY, Zhang YJ, Zalucki MP, Wu QJ and Yin XH, Population dynamics of the diamondback moth, *Plutella xylostella* (L.), in northern China: the effects of migration, cropping patterns and climate. *Pest Manag Sci* **74**(8): 1845-1853 (2018).
49. Peterson BK, Weber JN, Kay EH, Fisher HS and Hoekstra HE, Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**(5): e37135 (2012).
50. Li BY, Gao Q, Cao LJ, Hoffmann AA, Yang Q, Zhu JY and Wei SJ, Conserved profiles of digestion by double restriction endonucleases in insect genomes facilitate the design of ddRAD. *Zoological Systematics* **43**(4): 341-355 (2018).
51. Catchen J, Hohenlohe PA, Bassham S, Amores A and Cresko WA, Stacks: an analysis tool set for population genomics. *Mol Ecol* **22**(11): 3124-3140 (2013).
52. You MS, Yue Z, He WY, Yang XH, Yang G, Xie M, Zhan DL, Baxter SW, Vasseur L, Gurr GM, Douglas CJ, Bai JL, Wang P, Cui K, Huang SG, Li XC, Zhou Q, Wu ZY, Chen QL, Liu CH, Wang B, Li XJ, Xu XF, Lu CX, Hu M, Davey JW, Smith SM, Chen MS, Xia XF, Tang WQ, Ke FS, Zheng DD, Hu YL, Song FQ, You YC, Ma XL, Peng L, Zheng YK, Liang Y, Chen YQ, Yu LY, Zhang YN, Liu YY, Li GQ, Fang L, Li JX, Zhou X, Luo YD, Gou CY, Wang JY, Wang J, Yang HM and Wang J, A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics* **45**(2): 220-225 (2013).
53. Andrews S. FastQC: a quality control tool for high throughput sequence data (2010).
54. Catchen JM, Amores A, Hohenlohe P, Cresko W and Postlethwait JH, Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3* **1**(3): 171-182 (2011).

55. Langmead B and Salzberg SL, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**(4): 357 (2012).
56. Knaus BJ and Grünwald NJ, vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour* **17**(1): 44-53 (2017).
57. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R and Group GPA, The variant call format and VCFtools. *Bioinformatics* **27**(15): 2156-2158 (2011).
58. Rousset F, genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* **8**(1): 103-106 (2008).
59. Jombart T, Devillard S and Balloux F, Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics* **11**(94-108 (2010)).
60. Chen K-Y, Marschall EA, Sovic MG, Fries AC, Gibbs HL, Ludsin SA and Poisot T, *assignPOP*: An r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods Ecol Evol* **9**(2): 439-446 (2018).
61. Schmidt TL, van Rooyen AR, Chung J, Endersby-Harshman NM, Griffin PC, Sly A, Hoffmann AA and Weeks AR, Tracking genetic invasions: Genome-wide single nucleotide polymorphisms reveal the source of pyrethroid-resistant *Aedes aegypti* (yellow fever mosquito) incursions at international ports. *Evol Appl* **12**(6): 1136-1146 (2019).
62. Hardy OJ and Vekemans X, SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**(618-620 (2002)).
63. Kalinowski ST, Wagner AP and Taper ML, ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* **6**(2): 576-579 (2006).
64. Curtis J. *Farm insects*. Blackie and Son, London, (1860).
65. List G, Possible migration of diamondback moth. *J Econ Entomol* **30**(676 (1937)).
66. Wei SJ, Shi BC, Gong YJ, Jin GH, Chen XX and Meng XF, Genetic structure and demographic history reveal migration of the diamondback moth *Plutella xylostella* (Lepidoptera: Plutellidae) from the southern to northern regions of China. *PLoS One* **8**(4)2013).
67. Chu YI. The migration of diamondback moth. In *Diamondback moth management*. AVRDC, pp. 77-81 (1985).
68. Cao LJ, Wang ZH, Gong YJ, Zhu L, Hoffmann AA and Wei SJ, Low genetic diversity but strong population structure reflects multiple introductions of western flower thrips (Thysanoptera: Thripidae) into China followed by human-mediated spread. *Evol Appl* **10**(4): 391-401 (2017).
69. Otuka A, Migration of rice planthoppers and simulation techniques. *Planthoppers: new threats to the sustainability of intensive rice production systems in Asia*: 343-356 (2009).
70. Zhou L, Huang J and Xu H, Monitoring resistance of field populations of diamondback moth *Plutella xylostella* L.(Lepidoptera: Yponomeutidae) to five insecticides in South China: a ten-year case study. *Crop Protection* **30**(3): 272-278 (2011).

71. Chapman JW, Nilsson C, Lim KS, Bäckman J, Reynolds DR and Alerstam T, Adaptive strategies in nocturnally migrating insects and songbirds: contrasting responses to wind. *J Anim Ecol* **85**(1): 115-124 (2016).
72. Bretherton R. Lepidoptera immigration to the British Isles, 1969 to 1977. In *Proceedings and Transactions of the British Entomological and Natural History Society*, pp. 98-110 (1982).
73. Alerstam T, Hedenström A and Åkesson S, Long-distance migration: evolution and determinants. *Oikos* **103**(2): 247-260 (2003).
74. Mo JH, Baker G, Keller M and Roush R, Local dispersal of the diamondback moth (*Plutella xylostella* (L.)) (Lepidoptera : Plutellidae). *Environ Entomol* **32**(1): 71-79 (2003).
75. French RA, Long distance movement of two migrant Lepidoptera in relation to synoptic weather conditions. *Biometeorology* **2**(1): 565-569 (1967).
76. Xing K, Zhao F, Peng Y, Chang XQ and Ma CS, Typical migration pathways analysis of the diamondback moth *Plutella xylostella* in China in 2009. *Journal of Environmental Entomology* **38**(5): 896-902 (2016).
77. Hobson KA, Norris DR, Kardynal KJ and Yohannes E. Animal migration: a context for using new techniques and approaches. In *Tracking animal migration with stable isotopes*. Elsevier, pp. 1-23 (2019).
78. Van Doren BM and Horton KG, A continental system for forecasting bird migration. *Science* **361**(6407): 1115-1118 (2018).
79. Nilsson C, Dokter AM, Verlinden L, Shamoun-Baranes J, Schmid B, Desmet P, Bauer S, Chapman J, Alves JA, Stepanian PM, Sapir N, Wainwright C, Boos M, Górska A, Menz MHM, Rodrigues P, Leijnse H, Zehtindjiev P, Brabant R, Haase G, Weisshaupt N, Ciach M and Liechti F, Revealing patterns of nocturnal migration using the European weather radar network. *Ecography* **42**(5): 876-886 (2018).
80. Schmaljohann H, Radar aeroecology-a missing piece of the puzzle for studying the migration ecology of animals. *Ecography* **43**(2): 236-238 (2019).
81. Shamoun-Baranes J, Nilsson C, Bauer S and Chapman J, Taking radar aeroecology into the 21st century. *Ecography* **42**(5): 847-851 (2019).
82. Perry KD, Baker GJ, Powis KJ, Kent JK, Ward CM and Baxter SW, Cryptic *Plutella* species show deep divergence despite the capacity to hybridize. *BMC Evol Biol* **18**(1): 77 (2018).
83. Andrews KR, Good JM, Miller MR, Luikart G and Hohenlohe PA, Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**(2): 81-92 (2016).
84. Schmidt TL, Filipović I, Hoffmann AA and Rašić G, Fine-scale landscape genomics helps explain the slow spatial spread of *Wolbachia* through the *Aedes aegypti* population in Cairns, Australia. *Heredity* **120**(5): 386 (2018).
85. Uthicke S and Benzie JAH, Gene flow and population history in high dispersal marine invertebrates: mitochondrial DNA analysis of *Holothuria nobilis* (Echinodermata:

Holothuroidea) populations from the Indo-Pacific. *Mol Ecol* **12**(10): 2635-2648 (2003).

86. Manel S, Schwartz MK, Luikart G and Taberlet P, Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* **18**(4): 189-197 (2003).
87. Tabashnik BE, Liu YB, Malvar T, Heckel DG, Masson L, Ballester V, Granero F, Mensua JL and Ferre J, Global variation in the genetic and biochemical basis of diamondback moth resistance to *Bacillus thuringiensis*. *Proc Natl Acad Sci U S A* **94**(24): 12780-12785 (1997).
88. Baxter S, Zhao JZ, Gahan L, Shelton A, Tabashnik BE and Heckel D, Novel genetic basis of field-evolved resistance to Bt toxins in *Plutella xylostella*. *Insect Mol Biol* **14**(3): 327-334 (2005).

Figures legends

Fig. 1 Map of the collection sites of *Plutella xylostella* (red points) and inferred trajectories of migration in 2017 (a) and 2018 (b) based on assignment analysis of DAPC. The last number in the population code from 1-8 indicates January to August, respectively; population codes are described in Table 1. Ellipses represent different groups. The putative migration routes are represented by arrows. Arrow width represents the proportion of clusters from possible source population groups to recipient population groups estimated from the DAPC analysis (Fig. 2b and d).

Fig. 2 Analysis of population genetic structure of DBM using DAPC in 2017 (a, b) and 2018 (c, d). Figures a and c show scatter plots of DAPC analysis at the population level. Individuals from the same population are indicated by the same color. Figures b and d show the individual assignment table with one population per line. AIC values were plotted by the number of clusters in the bottom left panel of figures a and c. Populations collected in 2017 and 2018 were divided into four and five groups: ST, southern group; CE, central and eastern group; SW, southwestern group; YN, Yunnan group; SC, Sichuan group; NT, northern group.

Fig. 3 Membership probability of central and eastern (CE) and northern (NT) groups in 2017 (a) and 2018 (b). Assignment proportions of individuals were estimated in the R package *assignPOP*. Each individual was assigned to source groups ST and SW (2017) or ST, YN and SC (2018). Individuals in the CE group were from ST and SW (SC+YN in 2017), while individuals in the NT group were mainly from SW (SC+YN in 2018) group.

Fig. 4. Cousin networks superimposed on geographic coordinates of samples (a, e) and pairwise kinship values between related individuals plotted against geographical distances separating them (b-d, f-h). Results for populations of 2017 (a-d) and 2018 (e-h) are presented. Violin plots represent cousins identified by SPAGeDi analysis (b, f), the top 1% of Loiselle's *K* values (c, g) or the top 10% of Loiselle's *K* values (d, h) reflecting related individuals distributed across distance (km). In figures a and e, red lines indicate cousins

identified between southwestern (SW) group in 2017 or Yunnan (YN) group in 2018 and other group; blue lines indicate cousins between southern (ST) group and other groups; green lines indicate Sichuan (SC) group in 2018 and other groups; yellow lines indicate cousins within groups or between central (CE) and north (NT) groups. In figures b-d, f-h, population pairs were indicated for those with the largest geographic distance.

Table 1 Sampling information and population genetic diversity of *Plutella xylostella*. There was no private allele in any of the populations. All populations deviated from HWE with $p < 0.001$. A_R , allelic richness. %P, percentage of total alleles observed per locus per population. P_i , average nucleotide diversity. H_o , average observed heterozygosity per locus. H_e , average expected heterozygosity per locus. F_{IS} , inbreeding coefficient. CI, confidence interval. The first two letters in the population code indicate the province while the last two letters indicate the city of collection site.

Year	Group	Code	Collection location	Date	A_R	%P	P_i	H_o	H_e	F_{IS} (95% CI)	
2017	ST	HNHK	Hannan Province, Haikou	Mar. 28	1.39	81.40	0.097	0.070	0.089	0.208 (0.179-0.230)	
		GDGZ	Guangdong Province, Guangzhou	Mar. 30	1.40	82.19	0.103	0.075	0.092	0.185 (0.163-0.203)	
	CE	HNCS	Hunan Province, Changsha	Mar. 30	1.37	77.98	0.096	0.068	0.087	0.221 (0.178-0.256)	
		JXNC	Jiangxi Province, Nanchang	Apr. 24	1.40	82.17	0.100	0.072	0.091	0.204 (0.173-0.228)	
		HBXG	Hubei Province, Xiaogan	May 10	1.38	78.12	0.096	0.070	0.086	0.175 (0.135-0.207)	
	SW	HBWH	Hubei Province, Wuhan	May 10	1.40	81.00	0.099	0.072	0.090	0.192 (0.156-0.218)	
		AHHF	Anhui Province, Hefei	May 02	1.38	78.90	0.097	0.072	0.089	0.185 (0.160-0.206)	
		YNKM	Yunnan Province, Kunming	Jun. 04	1.39	79.57	0.100	0.070	0.090	0.212 (0.176-0.238)	
	NT	SCLS	Sichuan Province, Liangshan	May 12	1.35	74.62	0.096	0.070	0.085	0.165 (0.110-0.202)	
		CQCQ	Chongqing City	May 01	1.38	78.26	0.098	0.069	0.089	0.215 (0.178-0.241)	
		SCYA	Sichuan Province, Yaan	May 07	1.40	81.07	0.098	0.072	0.089	0.201 (0.171-0.223)	
		SCNC	Sichuan Province, Nanchong	May 09	1.37	78.24	0.094	0.067	0.083	0.195 (0.153-0.226)	
		SCCD	Sichuan Province, Chengdu	Apr. 14	1.37	78.88	0.096	0.070	0.086	0.187 (0.154-0.211)	
		SCGY	Sichuan Province, Guangyuan	May 09	1.37	77.90	0.095	0.069	0.086	0.202 (0.169-0.225)	
		BJYQ	Beijing City, Yanqing	May 05	1.39	81.07	0.096	0.071	0.086	0.178 (0.148-0.202)	
		HBZJ	Hebei Province, Zhangjiakou	May 21	1.37	75.62	0.099	0.073	0.088	0.164 (0.110-0.198)	
		2018	ST	HNHK	Hannan Province, Haikou	Mar. 06	1.34	72.57	0.094	0.067	0.083
GDSZ				Guangdong Province, Guangzhou	Apr. 21	1.39	79.62	0.099	0.076	0.088	0.152 (0.113-0.184)
GDGA	Guangdong Province, Guangzhou			Jan. 08	1.39	80.36	0.100	0.072	0.089	0.199 (0.166-0.227)	
GDGB	Guangdong Province, Guangzhou			Feb. 27	1.38	77.43	0.100	0.074	0.089	0.178 (0.126-0.216)	
YN	GDGC		Guangdong Province, Guangzhou	Apr. 03	1.39	78.40	0.098	0.071	0.088	0.187 (0.144-0.216)	
	GXNN		Guangxi Province, Nanning	Mar. 29	1.37	77.79	0.096	0.071	0.085	0.188 (0.146-0.220)	
	YNYX		Yunnan Province, Yuxi	Mar. 19	1.37	77.19	0.095	0.068	0.084	0.202 (0.148-0.241)	
	YNDH		Yunnan Province, Dehong	Apr. 10	1.32	70.31	0.092	0.070	0.082	0.142 (0.064-0.190)	
	YNKM		Yunnan Province, Kunming	Apr. 13	1.37	77.45	0.094	0.070	0.083	0.169 (0.127-0.196)	
	SCPZ		Sichuan Province, Panzhihua	Mar. 21	1.39	80.88	0.099	0.072	0.089	0.189 (0.151-0.221)	
	CE		FJFZ	Fujian Province, Fuzhou	Mar. 29	1.39	80.26	0.097	0.073	0.089	0.179 (0.141-0.208)
			HNCS	Hunan Province, Changsha	Apr. 11	1.40	81.60	0.100	0.073	0.090	0.180 (0.149-0.203)
JXNA			Jiangxi Province, Nanchang	Mar. 27	1.40	81.90	0.098	0.071	0.089	0.192 (0.161-0.220)	
JXNB			Jiangxi Province, Nanchang	Apr. 15	1.40	82.12	0.099	0.071	0.091	0.205 (0.177-0.226)	
JXNC			Jiangxi Province, Nanchang	May 23	1.39	80.86	0.099	0.071	0.088	0.193 (0.161-0.218)	
ZJNB			Zhejiang Province, Ningbo	Apr. 25	1.37	75.86	0.099	0.072	0.087	0.187 (0.135-0.221)	
SC	SCCD		Sichuan Province, Chengdu	Apr. 15	1.38	78.14	0.096	0.070	0.087	0.191 (0.149-0.221)	
	SCGY		Sichuan Province, Guangyuan	Apr. 15	1.38	79.74	0.096	0.069	0.087	0.197 (0.165-0.222)	
NT	NMAL		Neimenggu Province, Alashan	May 18	1.39	81.90	0.096	0.068	0.087	0.214 (0.180-0.243)	
	BJYA		Beijing City, Yanqing	Apr. 25	1.39	81.38	0.095	0.070	0.087	0.187 (0.155-0.212)	
	BJYB	Beijing City, Yanqing	May 02	1.41	82.60	0.097	0.073	0.088	0.169 (0.139-0.191)		
	BJYC	Beijing City, Yanqing	May 10	1.40	81.81	0.098	0.072	0.088	0.180 (0.149-0.204)		
	HBZA	Hebei Province, Zhangjiakou	Apr. 26	1.40	82.55	0.097	0.071	0.089	0.194 (0.166-0.216)		
	HBZB	Hebei Province, Zhangjiakou	Aug. 02	1.38	80.90	0.094	0.068	0.085	0.203 (0.169-0.229)		

Supplemental materials

Table S1 Summary statistics on the SNPs identified in populations of *Plutella xylostella* after filtering.

Table S2 Percentage assignment of individuals from the CE and NT groups to source groups.

Table S3 Percentage of related individuals among population groups estimated by SPAGEDi based on 1% and 10% of the most closely related individual pairs.

Fig. S1 Heat map of F_{ST} s among populations, the distribution of F_{ST} s and their averages for population pairs between and within groups.

Fig. S2 Scatter plots between pairwise genetic differentiation ($F_{ST}/(1-F_{ST})$) and geographical distances (ln km) among populations.

Fig. S3 Analysis of population genetic structure of DBM using DAPC across samples from the 2 years.

Fig. S4 Distributions of simulated kinship values based on SPAGeDi using R.

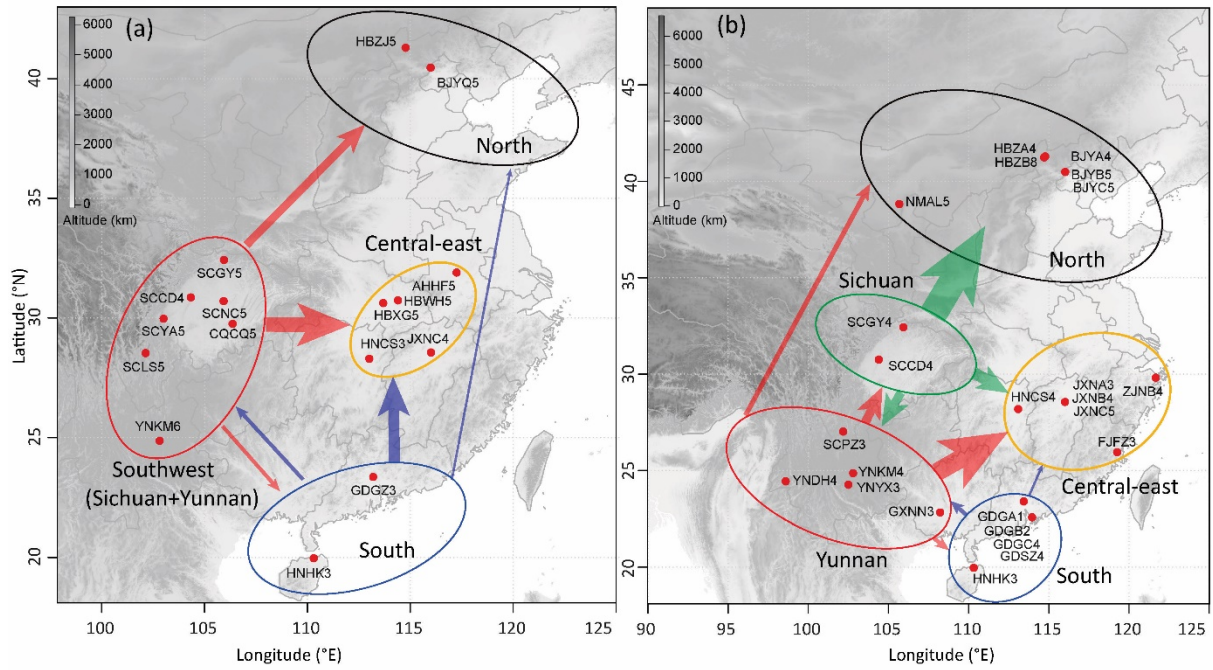


Fig. 1 Map of the collection sites of *Plutella xylostella* (red points) and inferred trajectories of migration in 2017 (a) and 2018 (b) based on assignment analysis of DAPC. The last number in the population code from 1-8 indicates January to August, respectively; population codes are described in Table 1. Ellipses represent different groups. The putative migration routes are represented by arrows. Arrow width represents the proportion of clusters from possible source population groups to recipient population groups estimated from the DAPC analysis (Fig. 2b and d).

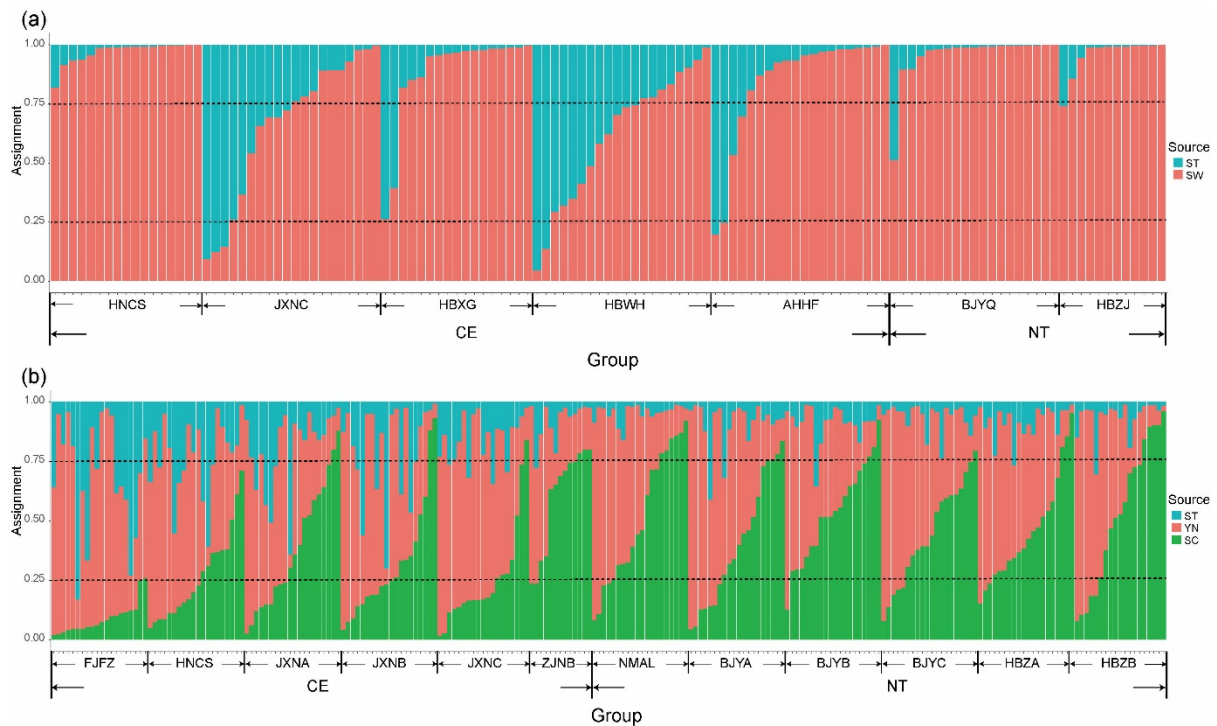


Fig. 3 Membership probability of central and eastern (CE) and northern (NT) groups in 2017 (a) and 2018 (b). Assignment proportions of individuals were estimated in the R package *assignPOP*. Each individual was assigned to source groups ST and SW (2017) or ST, YN and SC (2018). Individuals in the CE group were from ST and SW (SC+YN in 2017), while individuals in the NT group were mainly from SW (SC+YN in 2018) group.

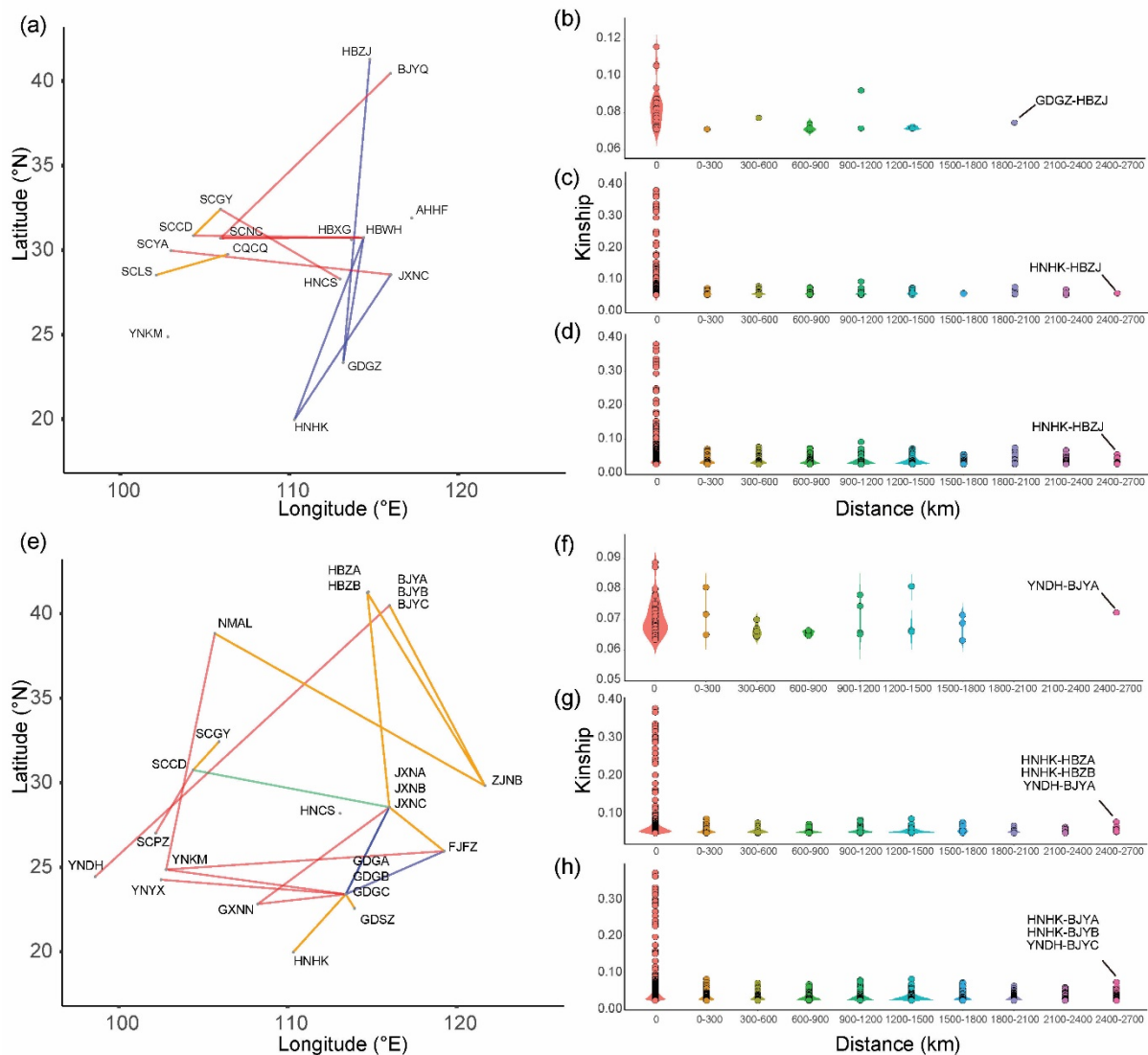
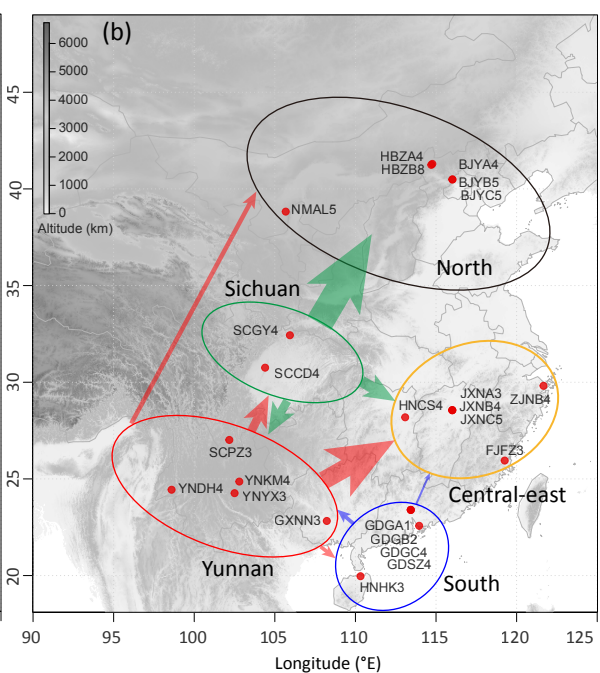
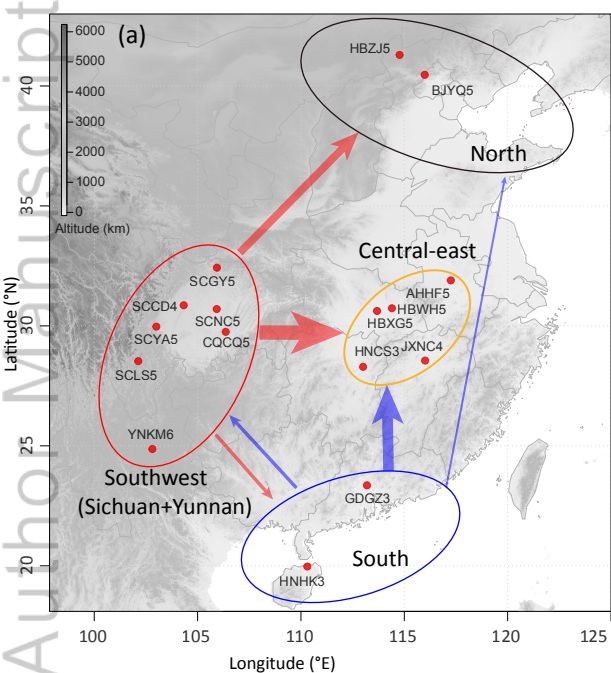
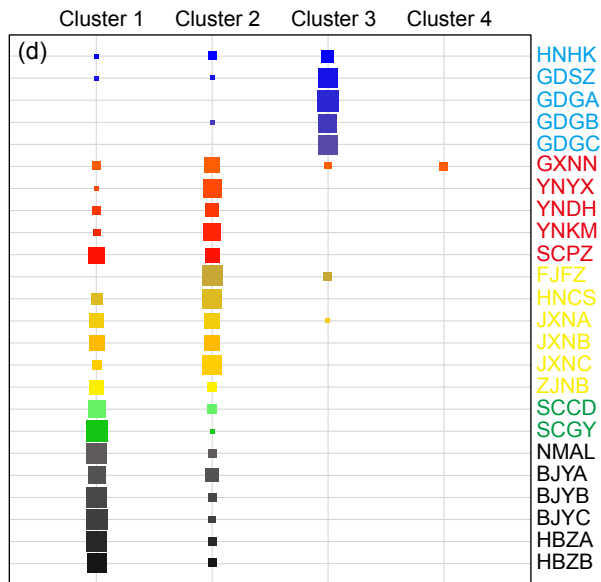
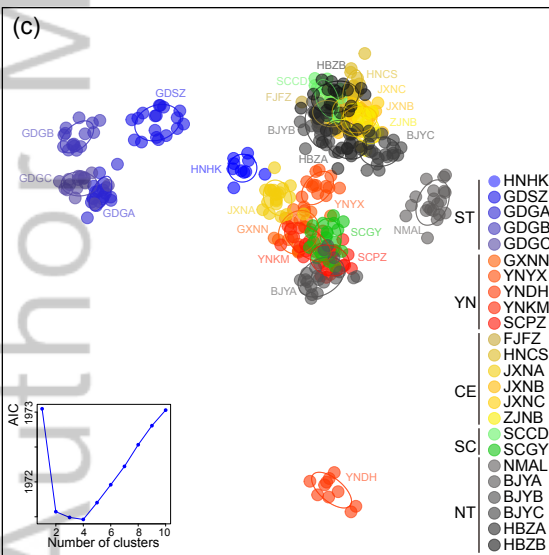
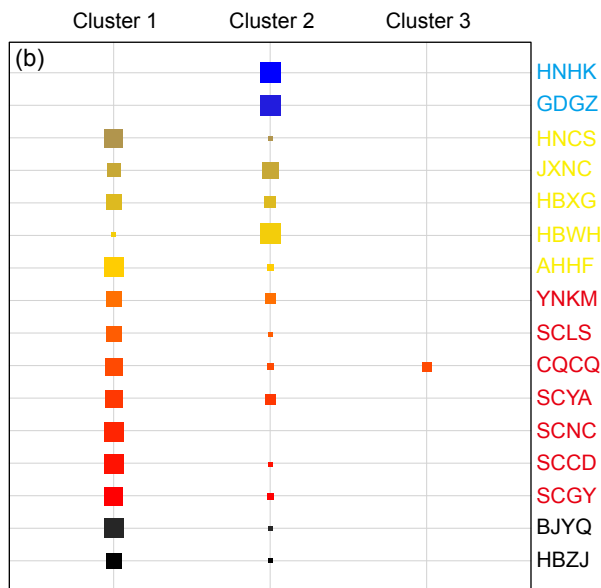
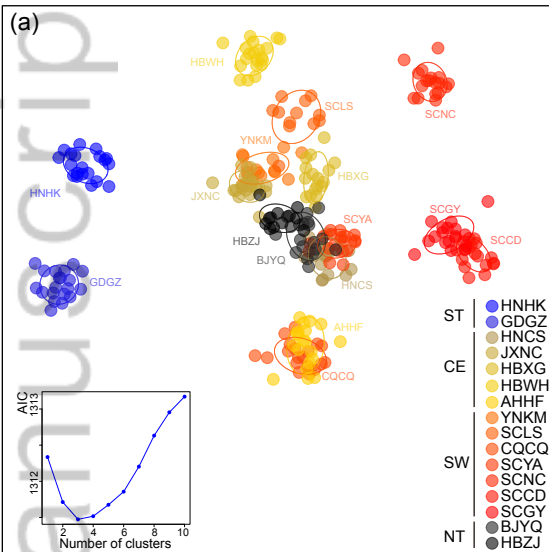
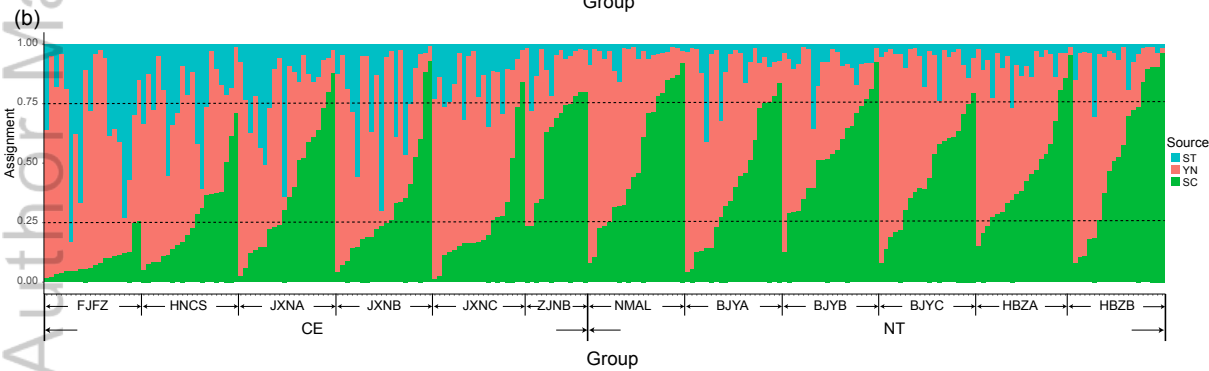
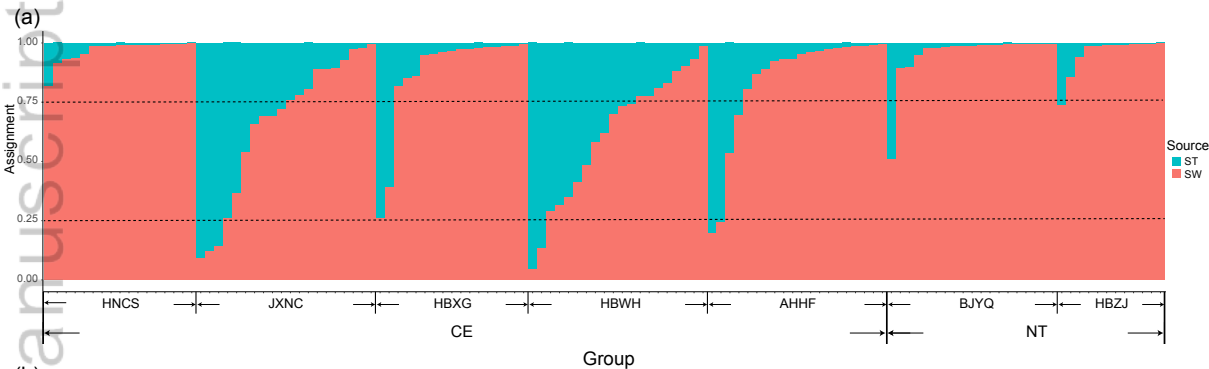
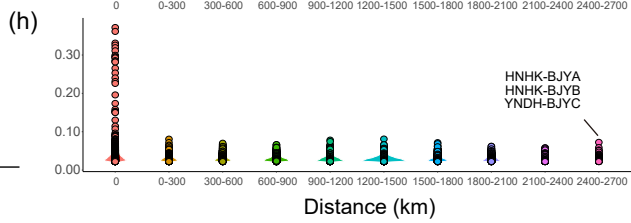
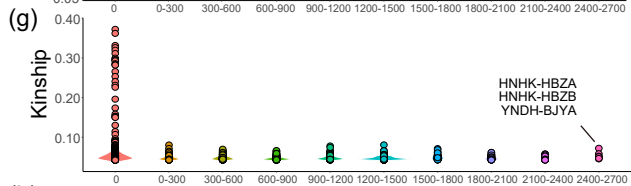
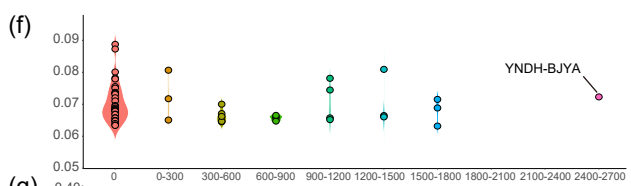
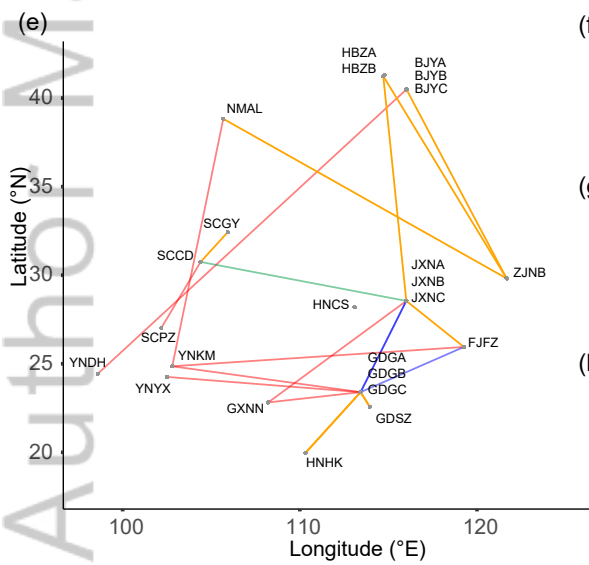
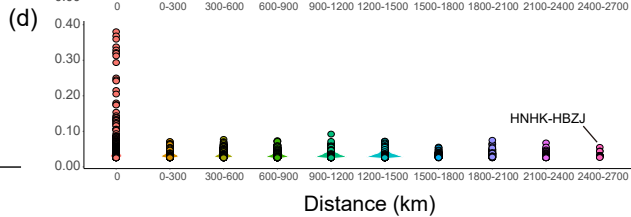
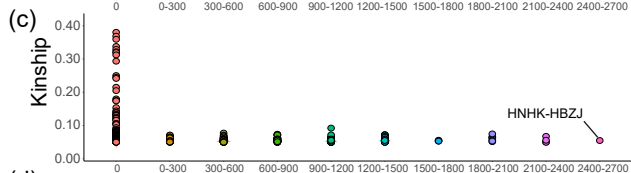
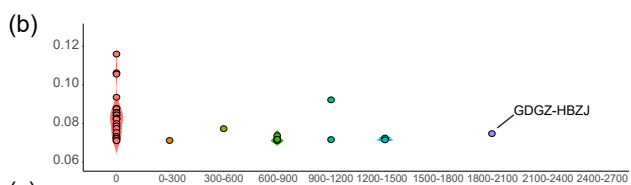
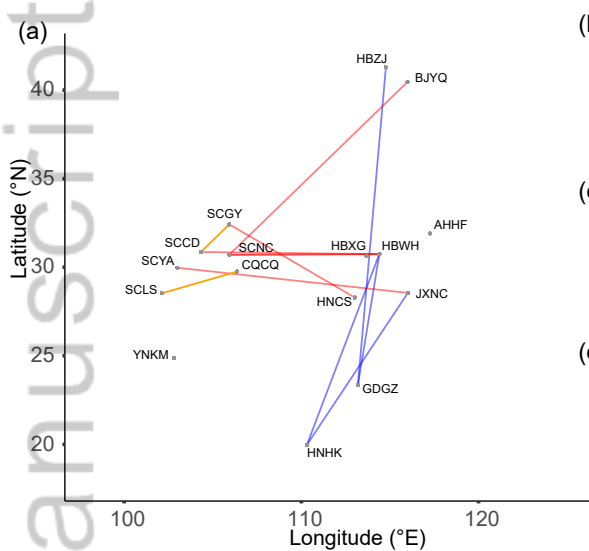


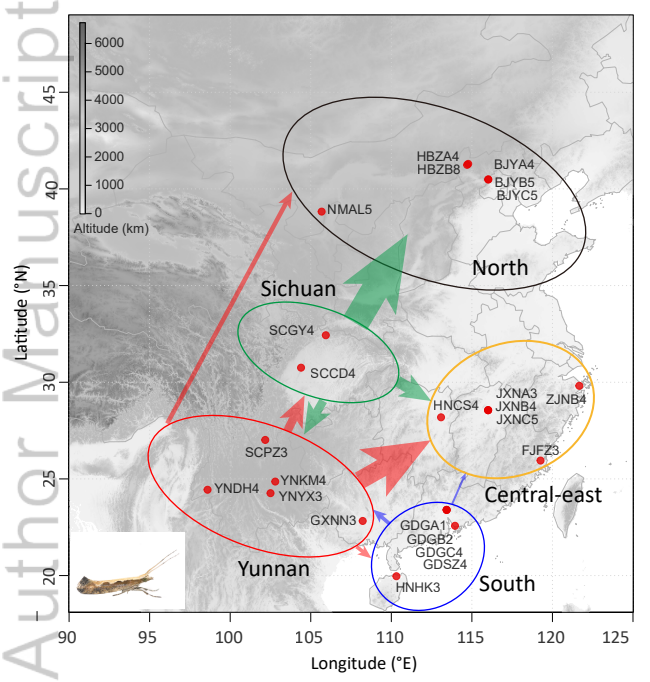
Fig. 4. Cousin networks superimposed on geographic coordinates of samples (a, e) and pairwise kinship values between related individuals plotted against geographical distances separating them (b-d, f-h). Results for populations of 2017 (a-d) and 2018 (e-h) are presented. Violin plots represent cousins identified by SPAGeDi analysis (b, f), the top 1% of Loiselle's *K* values (c, g) or the top 10% of Loiselle's *K* values (d, h) reflecting related individuals distributed across distance (km). In figures a and e, red lines indicate cousins identified between southwestern (SW) group in 2017 or Yunnan (YN) group in 2018 and other group; blue lines indicate cousins between southern (ST) group and other groups; green lines indicate Sichuan (SC) group in 2018 and other groups; yellow lines indicate cousins within groups or between central (CE) and north (NT) groups. In figures b-d, f-h, population pairs were indicated for those with the largest geographic distance.











Insects encompass the largest number of migratory animals, but their small body size impedes our understanding of insect migration in the wild, particularly the source populations of immigrants and distance of movement. This study unraveled migration trajectories of a pest insect, the diamondback moth, by deciphering connectivity among potential source and recipient populations from genomic variations.