MOLECULAR ECOLOGY
RESOURCES

Genetic structure and insecticide resistance characteristics of fall armyworm populations invading China

Journal:	*Molecular Ecology Resources*
Manuscript ID	MER-20-0047.R3
Manuscript Type:	Resource Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Zhang, Lei; Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences Liu, Bo Zheng, Weigang Liu, Conghui Zhang, Dandan Zhao, Shengyuan Li, Zaiyuan Xu, Peng jun Wilson, Ken Withers, Amy Jones, Christopher Smith, Judith Chipabika, Gilson Kachigamba, Donald Nam, Kiwoong d'Alençon, Emmanuelle Liu, Bei Liang, Xin-Yue Jin, Minghui Wu, Chao Chakrabarty, Swapan Yang, Xian-Ming Jiang, Yuying Liu, Jie Liu, Xiaolin Quan, Weipeng Wang, Guirong Fan, Wei Qian, Wanqiang Wu, Kongming Xiao, Yutao
Keywords:	Spodoptera frugiperda, subpopulations, population structure, resistance risk, gene insertion

Title: Genetic structure and insecticide resistance characteristics of fall armyworm populations invading China

Running title: Genome analysis of FAW invading China

Lei Zhang[1#], Bo Liu[1#], Weigang Zheng[1#], Conghui Liu[1#], Dandan Zhang[2#], Shengyuan Zhao[2#], Zaiyuan Li[1], Pengjun Xu[3,4], Kenneth Wilson[4,1], Amy Withers[4], Christopher M. Jones[5], Judith A. Smith[6], Gilson Chipabika[7], Donald L. Kachigamba[8], Kiwoong Nam[9], Emmanuelle d'Alençon[9], Bei Liu[1], Xinyue Liang[1], Minghui Jin[1], Chao Wu[1], Swapan Chakrabarty[1], Xianming Yang[2], Yuying Jiang[10], Jie Liu[10], Xiaolin Liu[11], Weipeng Quan[12], Guirong Wang[1], Wei Fan[1*], Wanqiang Qian[1*], Kongming Wu[2*], Yutao Xiao[1*]

1 Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China

2 State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China

3 Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao 266101, China

4 Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

5 Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre, Malawi

6 School of Forensic and Applied Sciences, University of Central Lancashire, Preston PR1 2HE, UK

7 Zambia Agriculture Research Institute (ZARI), Lusaka, Zambia

8 Department of Agricultural Research Services (DARS), Bvumbwe Research Station, P. O. Box 5748, Limbe, Malawi

9 DGIMI, Univ. Montpellier, INRA, Montpellier, 34095, France

10 National Agro-Tech Extension and Service Center, Beijing 100125, China

11 Novogene Bioinformatics Institute, Beijing 100083, China

12 Grandomics Biosciences, Co., Ltd, Beijing 102206, China

[#]These authors contributed equally to this work.

[*]Correspondence: Yutao Xiao (xiaoyutao@caas.cn), Kongming Wu (wukongming@caas.cn), Wanqiang

Qian (qianwanqiang@caas.cn), Wei Fan (fanwei@caas.cn).

Abstract

The rapid wide-scale spread of fall armyworm (*Spodoptera frugiperda*) has caused serious crop losses globally. However, differences in the genetic background of subpopulations and the mechanisms of rapid adaptation behind the invasion are still not well understood. Here we report the assembly of a 390.38Mb chromosome-level genome of fall armyworm derived from south-central Africa using Pacific Bioscience (PacBio) and Hi-C sequencing technologies, with scaffold N50 of 12.9 Mb and containing 22260 annotated protein-coding genes. Genome-wide resequencing of 103 samples and strain identification were conducted to reveal the genetic background of fall armyworm populations in China. Analysis of genes related to pesticide- and Bt-resistance showed that the risk of fall armyworm developing resistance to conventional pesticides is very high. Laboratory bioassay results showed that insects invading China carry resistance to organophosphate and pyrethroid pesticides, but are sensitive to genetically modified maize expressing the *Bacillus thuringiensis* (Bt) toxin Cry1Ab in field experiments. Additionally, two mitochondrial fragments were found to be inserted into the nuclear genome, with the insertion event occurring after the differentiation of the two strains. This study represents a valuable advance toward improving management strategies for fall armyworm.

KEYWORDS

Spodoptera frugiperda; subpopulations; population structure; resistance risk; gene insertion

1 INTRODUCTION

The fall armyworm, *Spodoptera frugiperda* (J.E. Smith), is a polyphagous pest that is native to tropical and subtropical America, with a strong capacity for migration and reproduction (Mitchell et al., 1991; Johnson, 1987; Westbrook et al., 2016). The pest was first detected in Africa in 2016 (Goergen et al., 2016) and spread to 44 African countries within two years. It was detected in India in 2018, and now it has spread to several southeastern Asian countries (Nagoshi et al., 2020). Such rapid spread poses a global threat to food production. The strong environmental adaptability of fall armyworm is not only reflected in its polyphagy for a wide range of host plants (Luginbill, 1928), but also in its evolution of resistance to chemical pesticides and genetically modified crops expressing *Bacillus thuringiensis* (Bt) toxins (Bernardi et al., 2015; Leibee & Capinera, 1995; Monnerat et al., 2015; Signorini et al., 2018; Storer et al., 2010). Studies have shown that gene families related to detoxification and metabolic processes in the fall armyworm have exhibited obvious expansion (Gouin et al., 2017; Liu et al., 2019). In addition, there are two morphologically identical, but genetically distinct, subpopulations or strains of fall armyworm, the rice-strain (R-strain) and the corn-strain (C-strain), which differ in their host plant selection and sex pheromone composition (Lima & McNeil, 2009; Pashley, 1986; Pashley, Hammond, & Hardy, 1992; Pashley & Martin, 1987). However, there is no absolute mating barrier between the two strains and productive hybridization has been confirmed in both laboratory and field studies (Dumas et al., 2015; Nagoshi, Meagher, Nuessly, & Hall, 2006).

To date, several field-evolved resistant populations of fall armyworm have been detected, including those displaying resistance to a variety of chemical pesticides and Bt crops (Chandrasena et al., 2018; Gutiérrez-Moreno et al., 2019; Zhu et al., 2015). The reported mechanisms of resistance to pesticides are

mainly due to variation in receptor genes, such as amino acid changes in the ryanodine receptor (RyR) (diamide), acetylcholinesterase (AChE) (organophosphate), voltagegated sodium channel (VGSC) (pyrethroids) (Boaventura et al., 2020; Carvalho et al., 2013; Yu, Nguyen, & Abo-Elghar, 2003). In addition, the frame-shift mutation resulting in early termination of the ATP-dependent Binding Cassette subfamily C2 gene (*ABCC2*) gene, caused by a 2-bp insertion, is linked to resistance to Bt toxin Cry1Fa (Banerjee et al., 2017). Field-evolved strains resistant to Bt toxin Vip3Aa20 were obtained by screening homozygous resistance loci in F_2 generations in the laboratory (Yang et al., 2018). Clarifying the development of pesticide- and Bt-resistance in fall armyworm would be helpful in providing scientific support for the commercialization of genetically modified crops and Bt biopesticides.

Recent studies have indicated that the molecular identification of the C- and R- strains of fall armyworm is dependent on which markers are used (Meagher & Gallo-Meagher, 2003; Nagoshi, 2012). The early molecular markers based on mitochondrial *Cytochrome Oxidase Subunit I* (*COI*) and Z-chromosome-linked *Tpi* genes failed to accurately assign the strain genetic background (Juárez et al., 2014; Nagoshi, 2019; Nagoshi, Goergen et al., 2019; Nagoshi et al., 2017). The dominant populations of fall armyworm invading Africa and Asia were speculated to be hybrid populations based on these two molecular markers (Zhang et al., 2019). In addition, an Africa-specific haplotype, different from those native to the Americas, was also reported in African and Chinese samples based on the *Tpi* gene (Liu et al., 2019; Nagoshi, Goergen et al., 2019), which makes strain identification and population genetic structure more complicated. Therefore, a genome-wide analysis of the genetic characteristics of invasive fall armyworm is becoming imperative. Although several versions of the fall armyworm genome have now been published (Gouin et al., 2017; Kakumani, Malhotra, Mukherjee, & Bhatnagar, 2014; Liu et al., 2019; Nam et al., 2019; Nandakumar, Ma,

& Khan, 2017), a high-quality genome assembly from a different geographical source is a valuable addition to the genomic resources for this species. Moreover, the different biological properties of the C and R strains and the debate about strain identification will benefit from further genomic support and explanation. Here we report a chromosome-level genome sequence of a male moth from an inbred fall armyworm strain, which derived from field populations collected in Zambia in 2017 and would be classed as C-strain by *COI* genotype but possessed an Africa-specific *Tpi* haplotype which differs from the Western Hemisphere (henceforth American) R-strain and C-strain. We also re-sequenced 103 fall armyworm samples from 16 Provinces in China, as well as four samples collected from two African countries (Zambia and Malawi). The genome-wide genetic backgrounds of the invading fall armyworm samples were compared, and the insecticide-resistance risk was assessed based on analysis of potential resistance-related genes. Comparative genomic analyses of these data help to reveal the resistance-related mechanisms and the population genetic characteristics of fall armyworm, which may facilitate its future management.

2 MATERIALS AND METHODS

2.1 Samples and sequencing for genome assembly

The fall armyworm samples were collected from maize fields in Lusaka, Zambia, in 2017 and reared to produce an inbred strain. One male moth, derived from seven successive generations of single-pair sib mating, was selected for genomic sequencing for the primary assembly dataset and all other individuals used in the Hi-C and RNAseq experiments were from the same inbred strain. DNA was extracted using QIAGEN® Genomic DNA kit (Cat#13323, QIAGEN) followed by purity assessment and quantification with NanoDrop™ One UV-Vis spectrophotometer (Thermo Fisher Scientific, USA) and Qubit® 3.0

Fluorometer (Invitrogen, USA), respectively. About 0.5μg gDNA was used as input to generate a PCR-free Illumina genomic library using the Truseq Nano DNA HT Sample preparation Kit (Illumina, USA), with 350bp insert size and this library was sequenced in 2 ×150bp format on the Illumina NovaSeq 6000. 5μg gDNA from the same individual was used as an input for ~20kb insert libraries (SMRTbell Template Prep Kit 1.0, Cat#100-259-100, PacBio) sequenced on the PacBio Sequel (Pacific Biosciences, USA). Two 3rd instar larvae were selected for Hi-C library construction, nuclear DNA was cross-linked *in situ*, extracted, and then digested with restriction enzyme (DpnII). Hi-C libraries were amplified by 12-14 cycles of PCR and sequenced on the Illumina NovaSeq 6000 platform with 2×150-bp reads. In addition, three 5th instar larvae, three pupae, three female moths and three male moths were used for RNA sequencing. Total RNA was extracted using the RNeasy Mini extraction kit (Qiagen), NanoPhotometer® spectrophotometer (IMPLEN, CA, USA) and Qubit® 2.0 Flurometer (Life Technologies, CA, USA) were used to check the purity and concentration of RNA, respectively. 1 μg total RNA per sample was used to make indexed cDNA libraries using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. The libraries had insert sizes of 250-300 bp were sequenced on the Illumina NovaSeq 6000 platform with 150bp paired-end output.

2.2 Genome assembly and correction

The raw PacBio reads longer than 5kb were assembled into contigs using the software wtdbg2 v2.4 with the parameters "-p 0 -k 15 -AS 2 -s 0.05 -L 5000" (Ruan & Li, 2019). Arrow v2.1.0 (https://github.com/PacificBiosciences/GenomicConsensus) was used to correct assembly errors after comparing contigs with PacBio reads using pbalign v0.4.1 (https://github.com/PacificBiosciences/pbalign). The Illumina raw reads were filtered by trimming the adapter and low-quality regions using clean_adapter

v1.1 with the parameter "-a Both-adapter -r 75 -s 12" and clean_lowqual v1.0 with the parameter "-e 0.001 -r 75" (https://github.com/fanagislab/ assembly_2ndGeneration/tree/master/clean_illumina). The filtered Illumina reads were aligned to the assembled contigs by BWA mem v0.7.17 (Li & Durbin, 2009), and single base errors in the contigs were corrected by Pilon v1.21 (Walker et al., 2014).

2.3 Genome estimation and evaluation

A distribution analysis of 17 k-mer frequencies was performed to estimate the genome size of fall armyworm. The filtered Illumina reads were used as input to construct k-mer frequencies by jellyfish (https://github.com/gmarcais/Jellyfish). And then, the genome size G=K_num/K_depth, where the K_num is the total number of K-mer, and K_depth is the frequency occurring more frequently than the others (Li et al., 2010). We used the arthropoda gene set (odb9) to assess the integrity of the genome by Benchmarking Universal Single-Copy Ortholog (BUSCO) v3.0.2 (Simao et al., 2015).

2.4 Chromsome assembly based on Hi-C data

The Hi-C sequencing raw reads were filtered to remove reads containing <5 bases of adaptor sequence; >50% of bases with phred quality value of <19; and <5% of unknown base (N). Then filtered reads were aligned to the assembled contigs using bowtie2 (v2.2.3; http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) (Langmead and Salzberg, 2012). Invalid read pairs were filtered using default settings by HiC-Pro (v2.7.8; https://github.com/nservant/HiC-Pro) (Servant et al., 2015). LACHESIS (https://github.com/shendurelab/LACHESIS) (Burton et al., 2013) was applied to cluster, order and orient contigs based on the agglomerative hierarchical clustering algorithm. For each chromosome cluster, the ordered contigs were oriented by building a weighted, directed acyclic graph (WDAG). The orientation of each contig in each chromosomal group was predicted according to the

maximum likelihood path through WDAG. Finally, we cut the chromosomes predicted by LACHESIS into bins of equal length (100kb) and constructed a heatmap based on the interaction signals revealed by valid mapped read pairs between bins using HiC-Pro.

2.5 Gene prediction and annotation

A *de novo* repeat library of fall armyworm was constructed by RepeatModeler v1.0.4 (http://www.repeatmasker.org/RepeatModeler.html). TEs were identified by RepeatMasker v4.0.6 (http://www.repeatmasker.org/) using both *de novo* library and Repbase library (Repbase-20150923), and tandem repeats were predicted using Tandem Repeats Finder (Benson, 1999) v4.07b. We used a combination of *ab initio* prediction, homology searches and RNA-seq annotation to predict genes in *Spodeptera frugiperda* genome. We performed the *ab initio* prediction using Augustus 2.5.5 with default parameters (Stanke and Waack, 2003). For homology-based annotation, we queried the *S. frugiperda* genome sequences against a database containing non-overlapping protein sequences from closed related species (*Bombyx mori*, *Helicoverpa armigera*, *Spodoptera litura*) by genBlastA with default parameters (She et al., 2009). Genewise (Birney et al., 2004) was used to refine the genBlastA mappings to the genome. For the RNA-seq annotation, the RNA-seq data were mapped to the assembled genome of *S. frugiperda* using Tophat v2.0.12 and alignments were processed by cufflinks v2.2.1 with default parameters to generate transcript predictions (Trapnell et al., 2012). EVidence Modeler (Haas et al., 2008) v1.1.1 was used to combine *ab initio* predictions, homology-based searches, and RNA sequencing alignments. Predicted gene models supported by at least one of the annotations using UniProt datbase, NR database, and RNA-seq data were retained. Gene functional annotation was performed by aligning the predicted protein sequences to the NCBI NR, UniProt, eggNOG, and KEGG databases with BLASTp v2.3.0+,

179 applying an E-value cut-off < 10^{-5}.

180 **2.6 Phylogenetic tree construction and genomic comparison**

181 Orthologous and paralogous gene families identified in a set of 10 species (*Drosophila melanogaster*,
182 *Plutella xylostella*, *Bombyx mori*, *Manduca sexta*, *Danaus plexippus*, *Heliconius melpomene*, *Operophtera*
183 *brumata*, *Helicoverpa armigera*, *Spodoptera frugiperda*, *Spodoptera litura*) with published genomes were
184 analyzed by OrthoFinder v2.3.1 with default parameters. Orthologous groups that contain single-copy
185 genes for each species were selected to construct the phylogenetic tree. The multi-sequence alignment of
186 proteins was accomplished by MUSCLE (Edgar, 2004) v3.8.31. A Neighbor-Joining (NJ) phylogenetic tree
187 was constructed using MEGA v7.0.14. The current assembled genome was aligned with two published
188 versions of fall armyworm genomes using MUMmer3.23 (Kurtz et al., 2004) package with cutoff of
189 identity >80% and coverage >80%. Alignments were filtered to generate a multi-alignment dataset using
190 the delta-filter utility with 85% minimum identity (-i 85) and minimum alignment length 10 (-l 10). A set of
191 unique alignments was created using the same filter criteria but with the addition of the -r and -q flags.

192 **2.7 Sampling for resequencing and population genetic study**

193 A total of 103 Chinese fall armyworm samples were used for resequencing. All samples were collected as
194 larvae on maize or sugarcane from 50 cities of 16 provinces (autonomous regions or municipalities) of
195 China. The larvae were fed with fresh maize leaves and brought back to the laboratory under ambient
196 conditions during transportation. Larval bodies were cleaned and then stored in a freezer at -80 °C. The
197 detailed sample information is shown in Supplementary Table 1 and the sample distribution in China is
198 shown in Supplementary Figure 1. In addition, four fall armyworm samples from Africa were also used for
199 resequencing, including two samples (AFR4-5) from the same inbred strain (AFR2017) as the genome

sequencing in this study, and another two samples (AFR14-15) which were collected from maize fields in Bvumbwe, Malawi, in January 2019, which is also an inbred strain (AFR2019) reared in laboratory. A total amount of 1.5μg gDNA of each sample was used to construct a 350-bp insert library using Truseq Nano DNA HT Sample preparation Kit (Illumina USA) sequenced in 150bp paired-end mode as described in section 2.1. Raw reads were aligned to the NCBI NT database using blastn, and reads with significant matches (identity >95% and coverage >80%) to microbes or host plants were removed.

A further 173 fall armyworm samples from 21 provinces in China were used for strain identification and molecular detection using PCR amplification and Sanger sequencing. Genomic DNA was extracted using Multisource Genomic DNA Miniprep Kit (Axygen, New York, USA) according to product instructions. The 50 μL PCR reaction mixture contained 25 μL of 2×Easytaq mix, a mixture of 2 μL forward and reverse primers (10 μmol/L), 2 μL of DNA, and 21 μL of Diethy1 pyrocarbonate (DEPC) H_2O. PCR was performed at 94°C for 5 min, 34 cycles of (94°C 30 s, 60°C 30 s, 72°C 30 s) and 72°C for 5 min. A total of 10 μL of PCR products containing the target fragment were sequenced by Life Technology (Shanghai, China). These samples were collected from the field as larvae or adult moths. Detailed sample information is shown in Supplementary Table 2 and the sample distribution in China is shown in Supplementary Figure 1. Mitochondrial *COI* and *Tpi* markers were used for strain identification. ABCC2 and AChE genes were detected based on primers designed according to published mutation sites (Banerjee et al., 2017; Carvalho et al., 2013). Inserted mitochondrial fragments in the nuclear genome were detected using primers designed in this study. All primer sequence information in this study is shown in Supplementary Table 3.

2.8 Read mapping and SNP calling

The Illumina raw reads from re-sequenced samples were filtered using clean_adapter and clean_lowqual

221 software as described in section 2.1, resulting in high-quality reads with an average error rate of <0.01.
222 Then, the high quality reads were aligned to the fall armyworm reference genome (American C-strain) and
223 mitochondrial genome sequences using BWA mem software (Li & Durbin, 2009) v0.7.5a with default
224 parameters. Alignments for each sample were processed by removing duplicate reads using SAMtools (Li,
225 Handsaker et al., 2009) software package v1.3. The mpileup function in SAMtools was used to generate
226 mpileup files for each sample. VCFtools (Li, 2011) was used to identify SNPs and small indels. Several
227 criteria were considered in SNP filtering: (1) a read mapping score higher than 40; (2) minimum coverage
228 greater than 10; (3) SNP genotypes called in >90% of samples. We also conducted principal component
229 analysis (PCA) to evaluate genetic structure using the software Genome-wide Complex Trait Analysis
230 (GCTA) version 1.04 (Yang, Lee, Goddard, & Visscher, 2011).

231 **2.9 Bioassays of insecticides and Bt maize in the field**

232 Bioassays were conducted by a topical application procedure (Armes, Jadhav, Bond, & King, 1992). Two
233 inbred strains (cdcc and cdyc) collected from Yunnan Province and reared for multiple generations in the
234 laboratory, were tested using 14 types of pesticide commonly used in agricultural production
235 (Supplementary Table 4). 1.0 µL drops of a serial dilution of technical insecticides in acetone solution were
236 applied with a micropipette to the thoracic dorsum of the 3rd instar larvae, with control larvae treated with
237 1.0 µL acetone. After treatment, the larvae were reared individually in 24-well plates containing ad libitum
238 artificial diet without any Bt proteins and insecticides. Larvae were retained in an insect chamber with a
239 controlled environment of 26 ± 1°C, 60 ± 10% RH and a photoperiod of 16 h: 8 h (L: D). Mortality was
240 assessed after 72 h treatment. Larvae were considered dead if they were unable to move in a coordinated
241 manner when prodded with a small soft brush. We used median lethal doses LC_{50} to evaluate the resistance

level of different fall armyworm populations. The LC_{50} and 95% fiducial limit (FL) for each insecticide were estimated by probit analysis using the software package POLO-PC (Russell, Robertson, & Savin, 1977) (LeOra Software, Berkeley, CA, USA).

The Bt toxin field bioassays were conducted at a genetically modified (GM) test base in Yunnan Province, China. Test seeds of GM maize (expressing Cry1Ab) and control maize were provided by DBN Biotech Center, Beijing DBN Technology Group Co., Ltd. Both maize types were planted in approximately 180 m^2, with each type being replicated three times. Larval density and maize damage rates were investigated at different growth stages of maize at seven different dates during June to July. The investigation was performed in a five-spot-sampling method with 20 maize plants per point. Fall armyworm damage assessment was performed according to standard procedures (Davis, Ng, & Williams, 1992; Williams, Buckley, & Daves, 2006; Wiseman & Widstrom, 1984).

3 RESULTS

3.1 High-quality genome assembly of fall armyworm

A total of 25.89 Gb raw PacBio long reads and 162.4 Gb Illumina raw reads were generated. After filtering low-quality and duplicated reads, 24.72 Gb PacBio long reads and 95.4 Gb high-quality Illumina reads were used for genome assembly, together representing an ~300 × coverage of the fall armyworm genome. Using wtdbg2 (Ruan & Li, 2019), the final genome was assembled into 776 contigs with size of 390.38 Mb and contig N50 length of 5.6 Mb (longest, 18.5 Mb), including a complete mitochondrial sequence (Table 1). The assembled genome size was close to the estimated size of 395 Mb based on k-mer depth distribution analysis, which was also similar to that of flow cytometry (396 ± 3Mb) (Gouin et al., 2017).

After interaction analysis based on a total of 78 Gb data obtained through Hi-C sequencing, 143 contigs were concatenated to 31 linkage groups with a scaffold N50 of 12.9 Mb, accounting for 96.3% of total genome length (Fig. 1). By aligning the Illumina data with the assembled fall armyworm genome, the mapping rate and coverage were 98.8% and 99.7% (≥5 reads) respectively, which showed the accuracy and high integrity of genome assembly. The genome size reported in this study is intermediate between those of previously published fall armyworm versions, but the genome is nearly 140 Mb smaller than that recently published by Liu et al. (2019). Genome collinearity analysis showed that more than 98% of the current assembled genome was consistent with previously published versions (Gouin et al., 2017; Liu et al., 2019) (Supplementary Table 5), and regions within the assembly presented in this study align to multiple regions of Liu's assembly, indicating the previous assembled genome with larger size was mainly caused by high heterozygosity of sequenced samples.

By combining homology-based and *de novo* approaches, we identified ~27.2% of repetitive elements in the assembled fall armyworm genome. Among the known repeat families, LINE constituted the most abundant repeat family, representing 8.7% of the repetitive sequences, while LTR was only 1.4% (Supplementary Table 6). To annotate the fall armyworm genome, we performed deep transcriptome sequencing of larvae, pupae, male and female moths, including three different developmental stages, which generated 98.4 Gb of RNA sequencing data. By combining homologue-based, *ab initio* and transcriptome-based approaches, we predicted 22260 protein-coding genes (gene models) in the fall armyworm genome, which is greater than the number of predicted genes in other lepidopteran genomes that have so far been published (Dasmahapatra et al., 2012; Kanost et al., 2016; Pearce et al., 2017; Wan et al., 2019; Xia et al., 2008; You et al., 2013; Zhan, Merlin, Boore, & Reppert, 2011). More than 85.5% of the predicted coding sequences

were supported by transcriptome sequencing data (defined as when ≥70% of the predicted coding sequence of a gene was covered by transcriptome reads). Further assessment of assembly integrity based on BUSCO analysis shows that the current genome contained 98.4% complete BUSCO genes.

Comparative analysis of orthogroups of nine Lepidoptera species and *Drosophila melanogaster* (Diptera) were performed (Supplementary Table 7). Among them, 17180 genes in 10755 orthogroups were found in the current genome of fall armyworm, and the remaining 5080 lineage-specific genes were identified as unassigned genes. Compared with *Spodoptera litura*, *S. frugiperda* has more species-specific genes, and the number of unassigned genes is much greater than that of *S. litura* (Fig. 2a). Phylogenomic analyses of the ten species were conducted using 1571 single-copy genes. As shown in Figure 2a, the taxonomic relationship and phylogenetic status of current species was similar to phylogenetic analyses based on 13 mitochondrial protein coding genes (Lämmermann, Vogel, & Traut, 2016). Three species of Noctuidae, including *S. frugiperda,* formed one group, which then clustered with *Bombyx mori* (Bombycidae) and *Manduca sexta* (Sphingidae). Two butterflies *Danaus plexippus* and *Heliconius melpomene* (both Nymphalidae) clustered together as an outer branch, while *Plutella xylostella* (Plutellidae) is the outermost branch of Lepidoptera (Fig. 2a).

3.2 Genetic background of fall armyworm population in China

A total of 103 fall armyworm samples from China were re-sequenced, as well as four samples from two countries in Africa (Zambia and Malawi). The generated Illumina data were ranged from 8.6Gb to 18.9Gb for each sample, with a median genome coverage of 32.5×. Firstly, we analyzed the whole mitochondrial genome sequences of all samples. A total of 208 SNP loci were selected for analysis, based on comparison of the published mitochondrial sequences of both the American R-strain (AXE) and C-strain (ASW)

(Gouin et al., 2017). Genotypes were obtained at these 208 sites for each individual after mapping the filtered sequence reads to the assembled mitochondrial genome. We found that most of the samples were assigned to R-strain, and all four samples from Africa were C-strain, while only four out of 103 samples in China were assigned to C-strain based on the mitochondrial genome (Fig. 3a). It should be noted that most R-strain samples surprisingly contain heterozygous mitochondrial SNPs, which could be caused by inserted C-strain fragments or existing standing variation of low frequency. The proportion of C-strain in this sample set was ~10% and was similar to that of the 173 Chinese fall armyworm samples identified by PCR based on the *COI* gene in this study (Supplementary Table 2).

Next, we analyzed the *Tpi* gene, which is commonly used in strain identification of fall armyworm (Nagoshi, 2012). By comparing the full length *Tpi* gene of American R-strain (AXE) and C-strain (ASW), 22 SNP loci were found. The genotype of each individual were analyzed based on these 22 sites. The results showed that all fall armyworm samples collected from China contained more C-strain SNP loci, as did the Malawi samples (AFR14, AFR15), but not those from Zambia samples (AFR4, AFR5) which represents the Africa-specific haplotype and which contained approximately 50% of R-strain SNP loci. Genotypes of seven Chinese samples were identical to the American C-strain (ASW) and the remainder samples contained a small proportion of R-strain genotypes or heterozygous SNPs (Fig. 3b). However, none of the samples was found to be identical to the American R-strain genotype (AXE). We further used PCR to analyze genotypes of 173 samples based on 10 strain-biased SNPs within the *Tpi* gene reported previously (Nagoshi, 2012). The results showed that almost all of the samples correspond to C-strain genotypes however three samples (G-GXW11, G-GXW13, G-EP6) were identified as Africa-specific haplotype, which was significantly different from known R- or C-strain genotypes (Fig. 4, Supplementary

Table 2). In summary, our genotyping results show that there are obvious contradictions between strain identification using mitochondrial and *Tpi* gene markers.

In order to clarify the genetic background of fall armyworm populations invading China, we screened a total of 707,353 SNPs exhibiting homozygous differences between the reference American R-strain (AXE) and C-strain (ASW) in the 107 re-sequenced samples (Fig. 3c). The results showed that all the samples, including the four from Africa, had more than 70% of the genetic background of the American C-strain (ASW) genotype. The proportion of R-strain SNPs was about 15%, and the remaining 12% were heterozygous. The results showed that fall armyworm invading China have a dominant percentage of the C-strain background. Principal component analysis (PCA) based on 5,998,089 whole-genome SNPs also demonstrated that samples from China were much closer to C-strain (ASW) than to R-strain (AXE), in which PC1 explained 6.45% of the variation. African samples from Zambia (AFR4, AFR5) were separated on PC2 which explained 2.15% of the variation (Fig. 3d). By comparing the results of the mitochondrial genome, *Tpi* gene and genome-wide identification, it becomes apparent that there is no correlation between the mitochondrial and whole genome genotype. Although *Tpi* genotyping shows results more similar to those of the whole genome, the presence of Africa-specific *Tpi*-haplotype increases the complexity of using this marker for identification.

3.3 Fall armyworm is developing high risk of resistance to conventional pesticides

Insecticide resistance evolution is one of the most challenging problems to be solved in the control of fall armyworm. Identifying resistance-related genes is helpful for the monitoring and prevention of fall armyworm outbreaks. We selected 14 previously reported resistance-related genes of lepidopteran pests and scanned the re-sequenced samples to analyze variation in target genes. The results showed that all the

target genes had multiple variation sites with high frequency of SNPs in the coding sequence (CDS) region (Supplementary Table 8).

Studies have shown that the amino acid substitutions in AChE (A201S, G227A, F290V), VGSC (T929I, L932F, L1014F) and RyR (I4790M, G4946E) result in resistance to organophosphate, pyrethroid and diamide insecticides, respectively. The results of variation scanning of the 107 re-sequenced samples showed that resistance mutations were found amino acids 201 and 290 of AChE (Fig. 5a). Among them, the first locus had 17.1% heterozygous mutations, and the third locus had 29.7% homozygous resistance mutations and 58.2% heterozygous mutations. No resistance mutations were detected at the targeted sites of the VGSC and RyR gene in any samples. We also designed primers to detect the resistance mutation sites in *AChE* in 173 Chinese samples by PCR amplification and Sanger sequencing. The results were similar to the Illumina data, showing that approximately 75% samples have homozygous or heterozygous variation at amino acid 290.

To understand the baseline susceptibility of fall armyworm invading China. We determined the LC_{50}s to 14 insecticides for two Chinese fall armyworm populations collected from Yunnan Province. The results showed that the LC_{50} for both fall armyworm populations to the fenvalerate, chlorpyrifos were at relatively high level, and well above those of the laboratory-susceptible *Helicoverpa armigera* strain (Bird & Downes, 2014). The LC_{50} to chlorantraniliprole were low along with emamectin benzoate and indoxacarb, which were similar to previous study on *H. armigera* and could be considered as susceptible baseline (Bird, 2015) (Fig. 6). The resistance levels of the two populations to pyrethroids and organophosphate pesticides were very high; in particular, the resistance ratios to chlorpyrifos of both populations were more than 300-fold compared to a laboratory susceptible fall armyworm population that was sampled in 1975 (Yu, 1991) (Fig.

5b). These results provide a susceptible baseline for fall armyworm populations invading China to different pesticides, which can provide guidance for resistance monitoring and pesticide management strategies.

3.4 Fall armyworm invading China are currently sensitive to Bt toxin in field-evolved experiment

The insertion of 2 bp in the *ABCC2* of fall armyworm was reported to cause a frame-shift mutation and results in resistance to Cry1Fa (Banerjee et al., 2017). We did not detect the same insertion mutation in 107 re-sequenced samples nor in 173 samples screened by using PCR and Sanger sequencing. Although the percentage of SNPs in the CDS region of other Bt receptors such as *SR-C* (scavenger receptor class C gene, a specific receptor for Vip3Aa in Sf9 cells), *TSPAN1* and other ABC gene-family related to Cry toxin were also very high (Supplementary Table 8), no reported resistant mutation were found in any target resistance genes.

The field tests showed that fall armyworm samples invading China were sensitive to genetically modified (GM) maize expressing Cry1Ab compared with control group. The damage assessment on larval density, percentage of damaged plants and average damage ratings of GM maize were significantly lower than those of the control group (Fig. 5c), which indicated that the GM maize expressing Cry1Ab currently has good control effects on the invading population of fall armyworm in China.

3.5 Insertion of mitochondrial fragments into nuclear genome in a recent evolution event

We found that two mitochondrial fragments, with sequence lengths of 1.5kb (partial *COI* gene and *NADH2* gene) and 1.6kb (partial *NADH*2 gene and 12S rRNA gene), were inserted into the nuclear genome, separated by a 4.1kb segment of the nuclear genome (Fig. 2b). The total length of a ~7.3kb fragment, including two inserted fragments, was supported by more than 28 raw reads of PacBio data. The lengths of all 28 reads were longer than 20kb and completely covered the 7.3kb fragment. However, the two

insertions were not found in other published fall armyworm genomes. In order to verify the accuracy of this result, we designed four primers based on flanking sequences of four connection points (Gap1-4 in Fig 2b), and the results of PCR amplification confirmed the existence of the insertion. The same primers were used in PCR assays to detect the insertion in 173 fall armyworm samples and it was found that the insertion was present in only 26.0% of all samples (Supplementary Table 2). At the same time, the resequencing data of 107 fall armyworm samples in this study also showed that there were varying numbers of reads covering the four junction points in 29 samples, and the percentage of samples with inserted reads was 27.1% (Supplementary Table 9). Both the PCR and resequencing results showed that the insertion was not present in all samples, perhaps suggesting that it has a recent evolutionary origin.

Moreover, the genotype of the two inserted mitochondrial fragments was identical to that of the C-strain, indicating that the insertion occurred after the differentiation of the R- and C-strains. Further analysis indicated that the two mitochondrial fragments were inserted into the intron region of lysine-specific demethylase 3 B (Kdm3B) gene, which is not likely to affect the expression of the gene. The inserted partial *COI* and *NADH2* gene fragments were also considered likely to be functionless.

DISCUSSION

The rapid spread of the fall armyworm has attracted popular attention worldwide. Accurate identification of its genetic characteristics (strain and pesticide resistance properties) has a direct and practical importance in terms of risk assessment and control strategies. A genome-wide analysis can reveal more in-depth genetic information than conventional gene-level analyses. The results of this study show that the fall armyworm invading China has a genetic background dominated by American corn-strain genotypes. Most of the fall

410 armyworm samples invading China were detected and collected from corn and sugarcane, which are more
411 likely to show the characteristics of C-strain host plants. Along the invasion path of the migratory fall
412 armyworm, there are large-scale rice planting areas in Southeast Asia and central China, however, there are
413 few reports of serious damage to rice caused by fall armyworm (http://www.fao.org/fall-armyworm). The
414 established R-strain fall armyworm in the Americas mainly feeds on turf grass, and there were few reports
415 of damage to rice in 1970's (Bowling, 1978; Gallego, 1967). In addition, the established R-strain *Tpi*
416 genotype has not been detected in any of the samples collected from Africa or Asia. So we speculate that
417 the American R-strain fall armyworm did not invade Africa or Asia, including China.
418 In our study, 103 re-sequenced Chinese samples were collected from different regions of 50 cities
419 distributed across 16 provinces (Supplementary Fig. 1). The collection time and sites coincided almost
420 perfectly with the spreading invasion of fall armyworm in China. However, there was no obvious
421 correlation between the time or site of collection and the genetic structure of the fall armyworm population
422 (Fig. 3). Almost all samples have similar genomic backgrounds, which suggests that the invading
423 population may originate from a single genetic source and there is no evidence for genomic selection
424 during the invasion.
425 According to our results, commonly used strain identification of fall armyworms by mitochondrial or *Tpi*
426 markers is limited or even inaccurate. The nuclear insertion of two C-strain partial *COI* fragments in this
427 study further underlines the need for caution in interpreting mitochondrial genotypes. We also found that
428 the AT/GC SNP located at *Tpi*-intron3 (P173/174) was inadequate as a diagnostic marker. In addition, the
429 TT/CC SNP located at *Tpi*-exon4 (P379/385) was associated with sequence variation in *Tpi*-intron4 (Fig. 4,
430 Supplementary Fig. 2), which could further be developed as markers to subdivide C-strain samples. It is

noteworthy that a special (Africa-specific) haplotype of the *Tpi* gene originally identified in Africa was tentatively designated as R-strain based on the $E4^{183}$ site (equal to P370 in Fig. 4 in this study) in previous studies (Nagoshi, 2012). Our genome-wide SNP analysis revealed that this haplotype contained more C-strain SNPs than R-strain.

The sample used for the genome sequencing in this study represents a combination of the special *Tpi* haplotype and C-strain *COI*. We also found combinations of the R-strain *COI* and special *Tpi* (sample G-XW13), as well as heterozygous forms of the *Tpi*-special and *Tpi*-C with the R-strain *COI* in two samples (G-GXW11, G-EP6). These combinations of different genotypes show that the genetic boundaries between two established (American) R- and C-strains are obscure. The insertion of two mitochondrial fragments into nuclear genome might be caused by random hybridization between different genotypes, which would suggest fall armyworm invading China might be descendants of an inter-strain hybrid population. This is the first report of DNA fragments transferred from mitochondria into the nuclear genome in *Spodoptera* lineage, and such two fragments could be used to develop markers to identify specific populations and to follow further evolutionary events of fall armyworm.

The rapid evolution of insecticide resistance and the increasing levels of resistance observed in fall armyworm populations needs attention. In this study, reported mutations related to insecticides resistance were detected in AChE gene. Although some mutation sites were detected as heterozygous in most samples at present, the frequency of resistant mutation sites will increase greatly under the selection pressure caused by application of related pesticides in field. The bioassay results showed that armyworms invading China have evolved high levels of resistance to organophosphate pesticides which was consistent with the results of molecular scanning of resistance-related genes, yet the resistance to pyrethroids pesticides cannot be

452 explained by reported mechanism. However, the fall armyworms invading China are currently sensitive to
453 GM maize expressing Cry1Ab in field experiments, and are also sensitive to other Bt toxins in the
454 laboratory, according to previous studies (Li et al., 2019). At present, GM maize shows better application
455 prospects in controlling fall armyworm in China, since larval density and damage rate of GM maize were
456 significantly less than that of non-GM plants, though this crop is currently not registered for use in the
457 country.
458 This study provides a high-quality reference genome that demonstrates a genomic feature different from the
459 established (American) C- or R-strain genotypes, as well as more comprehensive gene annotation. We also
460 present resequencing data for 103 fall armyworm individuals invading China. The samples cover different
461 regions and times during 2019, providing basic materials for analyzing global population genetic and
462 identifying patterns of invasiveness. Baseline resistance data for Chinese fall armyworm populations are
463 shown to 14 common pesticides, providing guidance for the control and resistance monitoring of fall
464 armyworm. Small-scale field experiments in this study suggest that fall armyworm in China are currently
465 susceptible to GM maize, and these results could provide an important application reference for
466 commercial planting of Bt maize in China. There are other important issues that remain for further
467 exploitation using this whole genome approach, such as identifying the genes involved in polyphagy,
468 migratory capability and olfaction, which could provide valuable tools for the future management of fall
469 armyworms.
470

471 # ACKNOWLEDGEMENTS

472 The following bodies provided funding that contributed to this work: Key Project for Breeding Genetic

Modified Organisms grant (2016ZX08012004-003, 2019ZX08012004-002), Shenzhen Science and Technology Program (KQTD20180411143628272), the UK's Global Challenges Research Fund and Biotechnology and Biological Sciences Research Council (BB/P023444/1), the UK Natural Environment Research Council Envision Doctoral Training Programme (NE/L002604/1), Key-Area Research and Development Program of Guangdong Province (2020B020223004).

REFERENCES

Armes, N. J., Jadhav, D. R., Bond, G. S., & King, A. B. (1992). Insecticide resistance in *Helicoverpa armigera* in South India. *Pesticide science*, 34(4), 355-364.

Banerjee, R., Hasler, J., Meagher, R., Nagoshi, R., Hietala, L., Huang, F., ... Jurat-Fuentes, J. L. (2017). Mechanism and DNA-based detection of field-evolved resistance to transgenic Bt corn in fall armyworm (*Spodoptera frugiperda*). *Scientific reports*, 7(1), 1-10.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic acids research*, 27(2), 573-580.

Bernardi, O., Bernardi, D., Ribeiro, R. S., Okuma, D. M., Salmeron, E., Fatoretto, J., ... Omoto, C. (2015). Frequency of resistance to Vip3Aa20 toxin from Bacillus thuringiensis in *Spodoptera frugiperda* (Lepidoptera: Noctuidae) populations in Brazil. *Crop Protection*, 76, 7-14.

Boaventura, D., Bolzan, A., Padovez, F. E., Okuma, D. M., Omoto, C., & Nauen, R. (2020). Detection of a ryanodine receptor target-site mutation in diamide insecticide resistant fall armyworm, *Spodoptera frugiperda*. *Pest management science*, 76(1), 47-54.

Bowling, C. C. (1978). Simulated insect damage to rice: effects of leaf removal. *Journal of Economic Entomology*, 71(2), 377-378.

Bird, L. J., & Downes, S. (2014). Toxicity and Cross-Resistance of Insecticides to Cry2Ab-Resistant and Cry2Ab-Susceptible *Helicoverpa armigera* and *Helicoverpa punctigera* (Lepidoptera: Noctuidae). *Journal of Economic Entomology*, 107(5), 1923-1930.

Bird, L. J. (2015). Baseline Susceptibility of *Helicoverpa armigera* (Lepidoptera: Noctuidae) to Indoxacarb, Emamectin Benzoate, and Chlorantraniliprole in Australia. *Journal of Economic Entomology*, 108(1), 294-300.

Birney, E., Clamp, M., & Durbin, R. (2004). GeneWise and genomewise. *Genome Res*, 14, 988-995.

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature biotechnology*, 31(12), 1119-1127.

Carvalho, R. A., Omoto, C., Field, L. M., Williamson, M. S., & Bass, C. (2013). Investigating the molecular mechanisms of organophosphate and pyrethroid resistance in the fall armyworm *Spodoptera frugiperda*. *PLoS ONE*, 8(4), e62268.

Chandrasena, D. I., Signorini, A. M., Abratti, G., Storer, N. P., Olaciregui, M. L., Alves, A. P., & Pilcher, C.

D. (2018). Characterization of field-evolved resistance to Bacillus thuringiensis-derived Cry1F δ-endotoxin in *Spodoptera frugiperda* populations from Argentina. *Pest management science*, 74(3), 746-754.

Dasmahapatra, K. K., Walters, J. R., Briscoe, A. D., Davey, J. W., Whibley, A., Nadeau, N. J., ... Salazar, C. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, 487(7405), 94-98.

Davis, F. M., Ng, S. S., & Williams, W. P. (1992). Visual rating scales for screening whorl-stage corn for resistance to fall armyworm. *Technical bulletin-Mississippi Agricultural and Forestry Experiment Station*(USA), 186, 1-9.

Dumas, P., Legeai, F., Lemaitre, C., Scaon, E., Orsucci, M., Labadie, K., ... Aury, J. M. (2015). *Spodoptera frugiperda* (Lepidoptera: Noctuidae) host-plant variants: two host strains or two distinct species?. *Genetica*, 143(3), 305-316.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.

Gallego, F. L. (1967). Lista preliminar de insectos de importancia económica y secundarios, que afectan los principales cultivos, animales domésticos y al hombre, en colombia. *Revista Facultad Nacional de Agronomía Medellín*, 26(65), 32-66.

Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A., & Tamò, M. (2016). First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera, Noctuidae), a new alien invasive pest in West and Central Africa. *PloS ONE*, 11(10), e0165632 (2016).

Gouin, A., Bretaudeau, A., Nam, K., Gimenez, S., Aury, J. M., Duvic, B., ... Kuwar, S. (2017). Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. *Scientific reports*, 7(1), 1-12.

Gutiérrez-Moreno, R., Mota-Sanchez, D., Blanco, C. A., Whalon, M. E., Terán-Santofimio, H., Rodriguez-Maciel, J. C., & DiFonzo, C. (2019). Field-evolved resistance of the fall armyworm (Lepidoptera: Noctuidae) to synthetic insecticides in Puerto Rico and Mexico. *Journal of economic entomology*, 112(2), 792-802.

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., ... Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome biology*, 9(1), 1-22.

Johnson, S. J. (1987). Migration and the life history strategy of the fall armyworm, *Spodoptera frugiperda* in the Western Hemisphere. *International Journal of Tropical Insect Science*, 8(4-5-6), 543-549.

Juárez, M. L., Schöfl, G., Vera, M. T., Vilardi, J. C., Murúa, M. G., Willink, E., ... Groot, A. T. (2014). Population structure of *Spodoptera frugiperda* maize and rice host forms in South America: are they host strains?. *Entomologia Experimentalis et Applicata*, 152(3), 182-199.

Kakumani, P. K., Malhotra, P., Mukherjee, S. K., & Bhatnagar, R. K. (2014). A draft genome assembly of the army worm, *Spodoptera frugiperda*. *Genomics*, 104(2), 134-143.

Kanost, M. R., Arrese, E. L., Cao, X., Chen, Y. R., Chellapilla, S., Goldsmith, M. R., ... Papanicolaou, A. (2016). Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, *Manduca sexta*. *Insect biochemistry and molecular biology*, 76, 118-147.

Kurtz S., Phillippy A., Delcher A.L, Smoot M., Shumway M., Antonescu C., & Salzberg S.L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, 5(2), 1-9.

Lämmermann, K., Vogel, H., & Traut, W. (2016). The mitochondrial genome of the Mediterranean flour

552 moth, *Ephestia kuehniella* (Lepidoptera: Pyralidae), and identification of invading mitochondrial
553 sequences (numts) in the W chromosome. *European Journal of Entomology*, 113, 482-488.
554 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4),
555 357-359.
556 Leibee, G. L. & Capinera, J. L. (1995). Pesticide resistance in Florida insects limits management
557 options. *Florida Entomologist*, 78, 386-399.
558 Li, G. P., Ji, T., Sun, X., Jiang, Y., Wu, K., & Feng, H. Q. (2019). Susceptibility evaluation of invaded
559 *Spodoptera frugiperda* population in Yunnan province to five Bt proteins. *Plant Protection*, 45(3), 15-20.
560 Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and
561 population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987-2993.
562 Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
563 transform. *Bioinformatics*, 25(14), 1754-1760.
564 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence
565 alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079.
566 Lima, E. R., & McNeil, J. N. (2009). Female sex pheromones in the host races and hybrids of the fall
567 armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Chemoecology*, 19(1), 29-36.
568 Liu, H., Lan, T., Fang, D., Gui, F., Wang, H., Guo, W., ... Sahu, S. K. (2019). Chromosome level draft
569 genomes of the fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae), an alien invasive pest
570 in China. *BioRxiv*, 671560.
571 Li, R., Fan, W., Tian, G., Zhu, H., He, L., Cai, J., ... & Wang, J. (2010). The sequence and de novo
572 assembly of the giant panda genome. *Nature*, 463(7279), 311-317.
573 Luginbill, P. (1928). The fall armyworm. U.S. Dept. Agric. Tech. Bull, 34, 1-91.
574 Meagher Jr, R. L., & Gallo-Meagher, M. (2003). Identifying host strains of fall armyworm (Lepidoptera:
575 Noctuidae) in Florida using mitochondrial markers. *Florida Entomologist*, 86(4), 450-455.
576 Mitchell, E. R., McNeil, J. N., Westbrook, J. K., Silvain, J. F., Lalanne-Cassou, B., Chalfant, R. B., ...
577 Proshold, F. I. (1991). Seasonal periodicity of fall armyworm, (Lepidoptera: Noctuidae) in the
578 Caribbean basin and northward to Canada. *Journal of entomological science*, 26(1), 39-50.
579 Monnerat, R., Martins, E., Macedo, C., Queiroz, P., Praca, L., Soares, C. M., ... Bravo, A. (2015). Evidence
580 of field-evolved resistance of *Spodoptera frugiperda* to Bt corn expressing Cry1F in Brazil that is still
581 sensitive to modified Bt toxins. *PLoS ONE*, 10(4), e0119544.
582 Nagoshi, R. N. (2012). Improvements in the identification of strains facilitate population studies of fall
583 armyworm subgroups. *Annals of the Entomological Society of America*, 105(2), 351-358.
584 Nagoshi, R. N. (2019). Evidence that a major subpopulation of fall armyworm found in the Western
585 Hemisphere is rare or absent in Africa, which may limit the range of crops at risk of infestation. *PloS*
586 *ONE*, 14(4), e0208966.
587 Nagoshi, R. N., Goergen, G., Du Plessis, H., van den Berg, J., & Meagher, R. (2019). Genetic comparisons
588 of fall armyworm populations from 11 countries spanning sub-Saharan Africa provide insights into
589 strain composition and migratory behaviors. *Scientific reports*, 9(1), 1-11.
590 Nagoshi, R. N., Htain, N. N., Boughton, D., Zhang, L., Xiao, Y., Nagoshi, B. Y., & Motasanchez, D.
591 (2020). Southeastern Asia fall armyworms are closely related to populations in Africa and India,
592 consistent with common origin and recent migration. *Scientific Reports*, 10(1), 1-10.
593 Nagoshi, R. N., Koffi, D., Agboka, K., Tounou, K. A., Banerjee, R., Jurat-Fuentes, J. L., & Meagher, R. L.
594 (2017). Comparative molecular analyses of invasive fall armyworm in Togo reveal strong similarities to

populations from the eastern United States and the Greater Antilles. *PLoS ONE*, 12(7), e0181982.

Nagoshi, R. N., Meagher, R. L., Nuessly, G., & Hall, D. G. (2006). Effects of fall armyworm (Lepidoptera: Noctuidae) interstrain mating in wild populations. *Environmental entomology*, 35(2), 561-568.

Nam, K., Gimenez, S., Hilliou, F., Blanco, C. A., Hänniger, S., Bretaudeau, A., ... d'Alençon, E. (2019). Adaptation by copy number variation increases insecticide resistance in fall armyworms. *BioRxiv*. 812958.

Nandakumar, S., Ma, H., & Khan, A. S. (2017). Whole-genome sequence of the *Spodoptera frugiperda* Sf9 insect cell line. *Genome Announc.*, 5(34), e00829-17.

Pashley, D. P. (1986). Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): a sibling species complex?. *Annals of the Entomological Society of America*, 79(6), 898-904.

Pashley, D. P., Hammond, A. M., & Hardy, T. N. (1992). Reproductive isolating mechanisms in fall armyworm host strains (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America*, 85(4), 400-405.

Pashley, D. P., & Martin, J. A. (1987). Reproductive incompatibility between host strains of the fall armyworm (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America*, 80(6), 731-733.

Pearce, S. L., Clarke, D. F., East, P. D., Elfekih, S., Gordon, K. H. J., Jermiin, L. S., ... Rane, R. V. (2017). Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC biology*, 15(1), 63.

Ruan, J., & Li, H. (2019). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, 1-4.

Russell, R. M., Robertson, J. L., & Savin, N. E. (1977). POLO: a new computer program for probit analysis. *Bulletin of the ESA*, 23(3), 209-213.

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annota- tion completeness with single-copy orthologs. *Bioinformatics*, 31, 3210–3212.

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C., Vert, J., ... & Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology*, 16(1), 259-259.

She, R., Chu, J. S. C., Wang, K., Pei, J., & Chen, N. S. (2009). genBlastA: Enabling BLAST to identify homologous gene sequences. *Genome Res*, 19, 143-149.

Signorini, A. M., Abratti, G., Grimi, D., Machado, M., Bunge, F. F., Parody, B., ... Araujo, M. P. (2018). Management of field-evolved resistance to Bt maize in Argentina: a multi-institutional approach. *Frontiers in bioengineering and biotechnology*, 6, 67, 1-5.

Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, 19, 215-225.

Storer, N. P., Babcock, J. M., Schlenz, M., Meade, T., Thompson, G. D., Bing, J. W., & Huckaba, R. M. (2010). Discovery and characterization of field resistance to Bt maize: *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Puerto Rico. *Journal of economic entomology*, 103(4), 1031-1038.

Trapnell, C., Roberts, A., Goff, L. A., Pertea, G., Kim, D., Kelley, D. R., ... & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, 7(3), 562-578.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., … Earl, A. M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly im- provement. *PLoS ONE*, 9, e112963.

Wan, F., Yin, C., Tang, R., Chen, M., Wu, Q., Huang, C., ... Wang, G. (2019). A chromosome-level

genome assembly of *Cydia pomonella* provides insights into chemical ecology and insecticide resistance. *Nature communications*, 10(1), 1-14.

Westbrook, J. K., Nagoshi, R. N., Meagher, R. L., Fleischer, S. J., & Jairam, S. (2016). Modeling seasonal migration of fall armyworm moths. *International journal of biometeorology*, 60(2), 255-267.

Williams, W. P., Buckley, P. M., & Daves, C. A. (2006). Identifying resistance in corn to southwestern corn borer (Lepidoptera: Crambidae), fall armyworm (Lepidoptera: Noctuidae), and corn earworm (Lepidoptera: Noctuidae). *J. Agr. Urban Entomol*, 23, 87-95.

Wiseman, B. R., & Widstrom, N. W. (1984). Fall armyworm damage ratings on corn at various infestation levels and plant development stages. *J. Agric. Entomol*, 1(2), 115-119.

Xia, Q. Y., Wang, J., Zhou, Z. Y., Li, R. Q., Fan, W., Cheng, D. J., … Xiang, Z. H. (2008). The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect biochemistry and molecular biology*, 38(12), 1036-1045.

Yang, F., Morsello, S., Head, G. P., Sansone, C., Huang, F., Gilreath, R. T., & Kerns, D. L. (2018). F2 screen, inheritance and cross-resistance of field-derived Vip3A resistance in *Spodoptera frugiperda* (Lepidoptera: Noctuidae) collected from Louisiana, USA. *Pest management science*, 74(8), 1769-1778.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*, 88(1): 76-82.

You, M., Yue, Z., He, W., Yang, X., Yang, G., Xie, M., ... Douglas, C. J. (2013). A heterozygous moth genome provides insights into herbivory and detoxification. *Nature genetics*, 45(2), 220-227.

Yu, S. J. (1991). Insecticide resistance in the fall armyworm, *Spodoptera frugiperda* (J. E. Smith). *Pesticide Biochemistry and Physiology*, 39(1), 84-91.

Yu, S. J., Nguyen, S. N., & Abo-Elghar, G. E. (2003). Biochemical characteristics of insecticide resistance in the fall armyworm, *Spodoptera frugiperda* (J. E. Smith). *Pesticide Biochemistry and Physiology*, 77(1), 1-11.

Zhang, L., Liu, B., Jiang, Y., Liu, J., Wu, K., & Xiao, Y. (2019). Molecular characterization analysis of fall armyworm populations in China. *Plant Protection*, 45, 10-27.

Zhan, S., Merlin, C., Boore, J. L., & Reppert, S. M. (2011). The monarch butterfly genome yields insights into long-distance migration. *Cell*, 147(5), 1171-1185.

Zhu, Y. C., Blanco, C. A., Portilla, M., Adamczyk, J., Luttrell, R., & Huang, F. (2015). Evidence of multiple/cross resistance to Bt and organophosphate insecticides in Puerto Rico population of the fall armyworm, *Spodoptera frugiperda*. *Pesticide Biochemistry and Physiology*, 122, 15-21.

DATA AVAILABILITY STATEMENT

This Whole Genome Shotgun project of *Spodoptera frugiperda* has been deposited at DDBJ/ENA/GenBank under the accession WUTJ00000000 with BioProject ID PRJNA591441. The version described in this paper is version WUTJ01000000. Raw sequencing reads of PacBio, RNA-seq, Hi-C, and re-sequencing in this paper can be accessed at ftp://ftp.agis.org.cn/Spodoptera_Frugiperda/.

675

676 # AUTHORS' CONTRIBUTIONS

677 Y. X., K. W., W. Q. and W. F. conceived the project, designed content and managed the project; L. Z. and
678 G. W. coordinated the project; B. L., Z. Li. and X. Liu. performed reads mapping, SNP calling and
679 population analysis; W. Z. and C. L. performed genome assembly and annotation; L. Z. and B. L.
680 performed Hic assembly; D. Z. performed the laboratory bioassay; S. Z. performed the field experiment; P.
681 X. performed transcriptome analysis; K. N. and E. A. provided the raw data of America R-strain and
682 C-strain; B. Liu., X. L., M. J., C. W. and X. Y. performed the DNA extraction, PCR and sequence variation
683 analysis; W. Q. constructed DNA libraries, performed sequencing; Y. J. and J. L. collected and provided
684 samples from China; L. Z. wrote the manuscript; K. W., A. W., C. M. J., J. A. S., G. C., D. L. K. and S. C.
685 revised the manuscript. All authors commented on the manuscript.

686

687

688 **TABLE 1** Summary of assembly results of *Spodoptera frugiperda*

Assembly feature	FAW (This study)	FAW (corn strain)	FAW (rice strain)
Assembled sequences (Mb)	390	438	371
Longest scaffold size (kb)	21916.7	943.2	314.1
N50 size of scaffold (kb)	12966.7	52.8	28.5
N90 size of scaffold (kb)	7574.2	3.5	6.4
Longest contig size (kb)	18555.4	362.9	191.4
N50 size of contig (kb)	5606.9	16.9	24.3
N90 size of contig (kb)	991.8	2.9	5.6
GC content in genome (%)	36.4	36.0	36.1
Number of gene models	22,260	21,700	26,329
BUSCO complete gene (%)	98.4	88.1	93.5
BUSCO duplicated gene (%)	2	11.3	2
BUSCO missing gene (%)	1.4	4.2	2.3

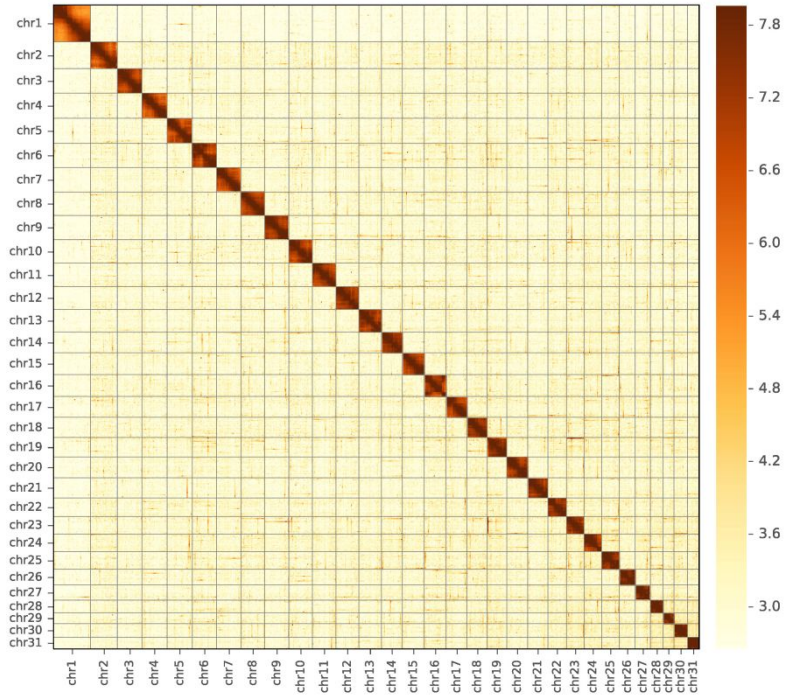


FIGURE 1 A genome-wide contact matrix from Hi-C data between each pair of the 31 chromosomes.

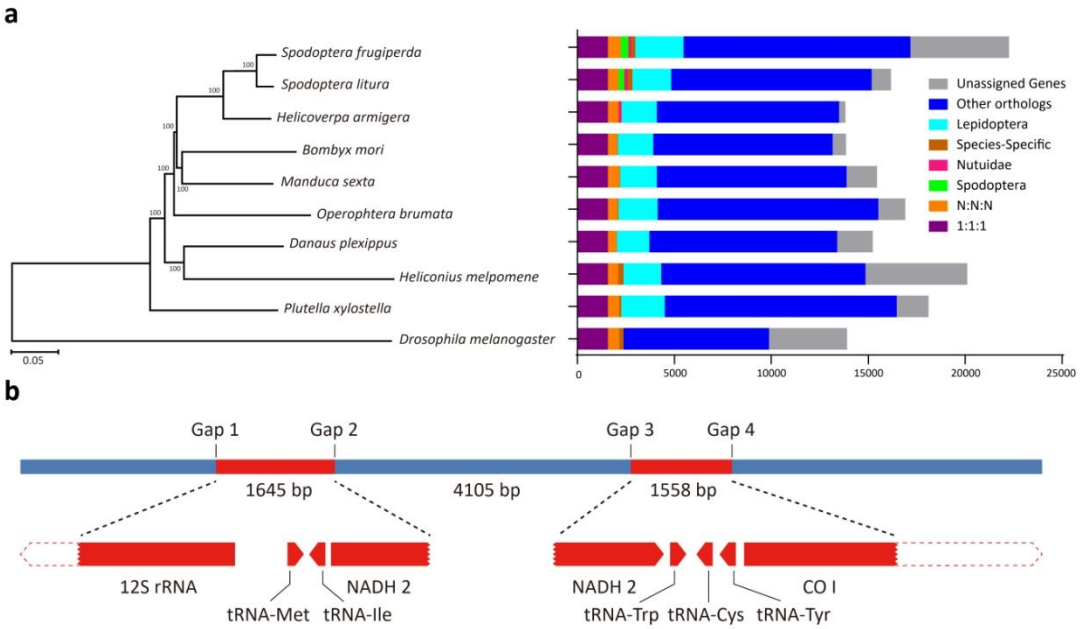


FIGURE 2 Phylogenetic relationships and schematic map of mitochondrial insertion. a) Phylogenetic tree and genomic comparison of 10 species of Lepidoptera and Diptera. *Drosophila melanogaster* was used as an outgroup and bootstrap value was set as 1000, 1:1:1 include the common orthologs with the same number of copies in different species, N:N:N include the common orthologs with different copy numbers in different species, other orthologs include the unclassified orthologs, and unassigned genes include the genes that cannot be clustered into known gene families. b) A schematic map of two mitochondrial fragments inserted into the nuclear genome, *NADH2* gene was separated by a 4105-bp fragment, and both two inserted mitochondrial fragments were identical with C-strain genotype.

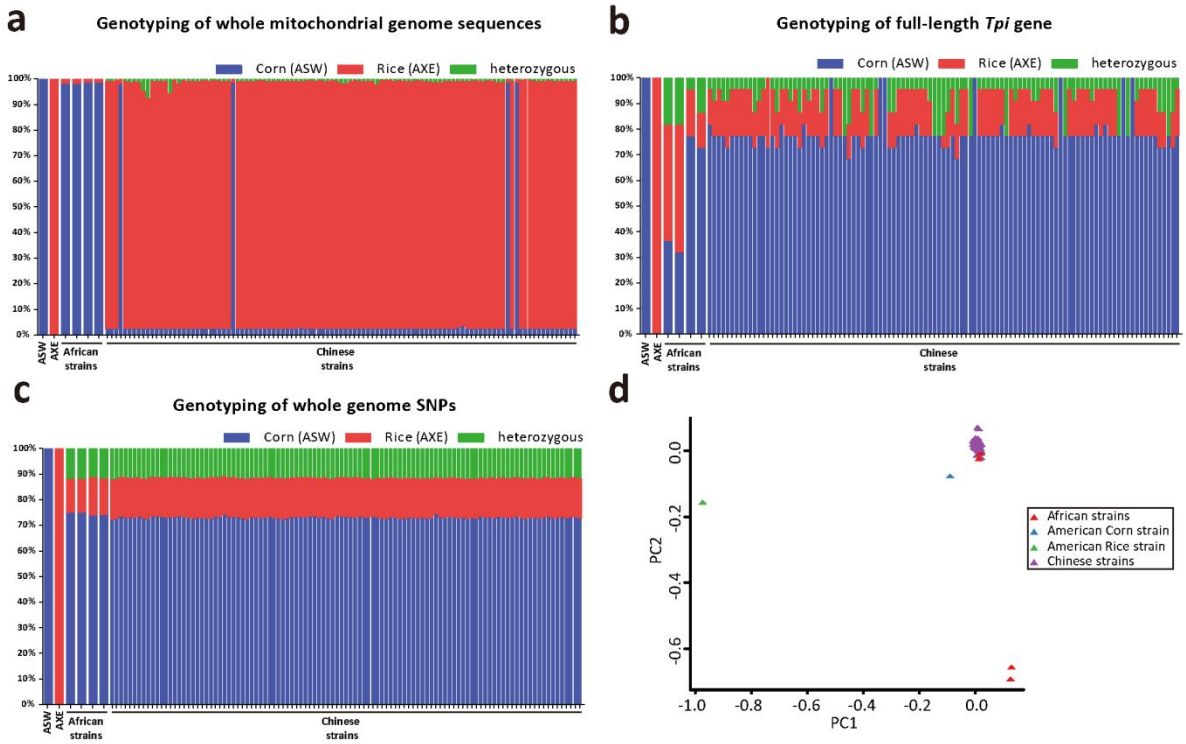


FIGURE 3 Genetic background of 107 fall armyworm samples. a) Genotyping based on 208 mitochondrial SNP loci. From left to right, the leftmost two samples were ASW (the American Corn strain) and AXE (the American Rice strain); four African strains (AFR4-5 from Zambia, then AFR14-15 from Malawi); then 103 strains from China and the order of each sample is consistent with Supplementary Table 1. b) Genotyping based on 22 SNP loci in the *Tpi* gene. c) Genotyping based on 707,353 genome SNP loci. d) Principal Component Analysis (PCA) based on 5,998,089 whole-genome SNPs. Colour codes indicate samples from different sources, the two samples at the bottom in red are African samples AFR4 and AFR5.

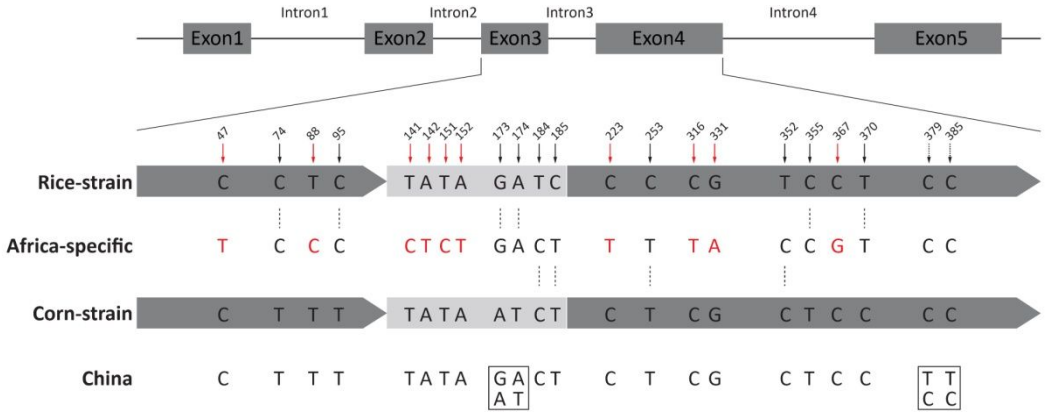


FIGURE 4 Diagram of the *Tpi* gene segments with respect to consensus Western Hemisphere sequences and the haplotypes observed in samples collected from Africa and China. Black solid arrows indicate 10 SNPs used to identify American R-strain and C-strain fall armyworm, in which P370 was considered to be an effective diagnostic marker especially. Red solid arrows indicate 10 SNPs specific to Africa-specific strain. The boxes represent two variable loci in some Chinese samples, including homozygous or heterozygous genotypes.

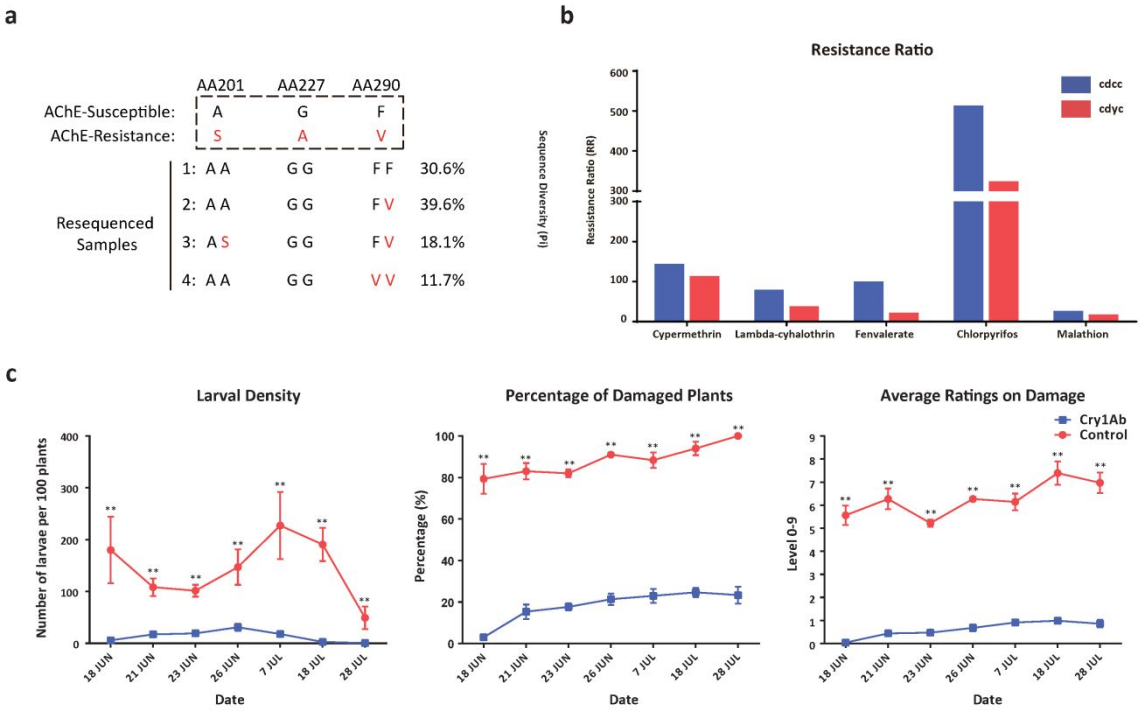


FIGURE 5 Genome scans and bioassays of fall armyworm for insecticides resistance. a) Genotype and resistance mutation sites of AChE gene in fall armyworm populations in China. b) The resistance ratios (RRs) of two Chinese fall armyworm populations to pyrethroids (cypermethrin, lambda-cyhalothrin, fenvalerate) and organophosphates (chlorpyrifos, malathion) insecticides, cdcc and cdyc represent two inbred strains collected from Yunnan Province in China. The RRs were calculated by LD_{50} (μg/g) of field population over the LD_{50} of a susceptible population as in Yu et al. (1991). c) Resistance tests of GM maize and non-GM maize to fall armyworm in field experiments. Error bar shows the SD (n=15), asterisk indicate significant differences base on Student's t-test (**p <0.01).

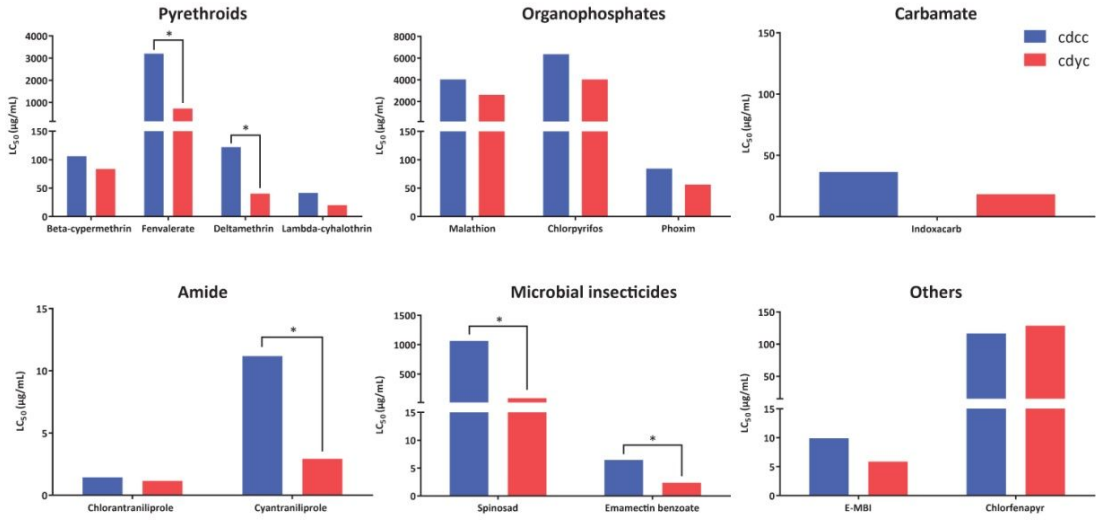
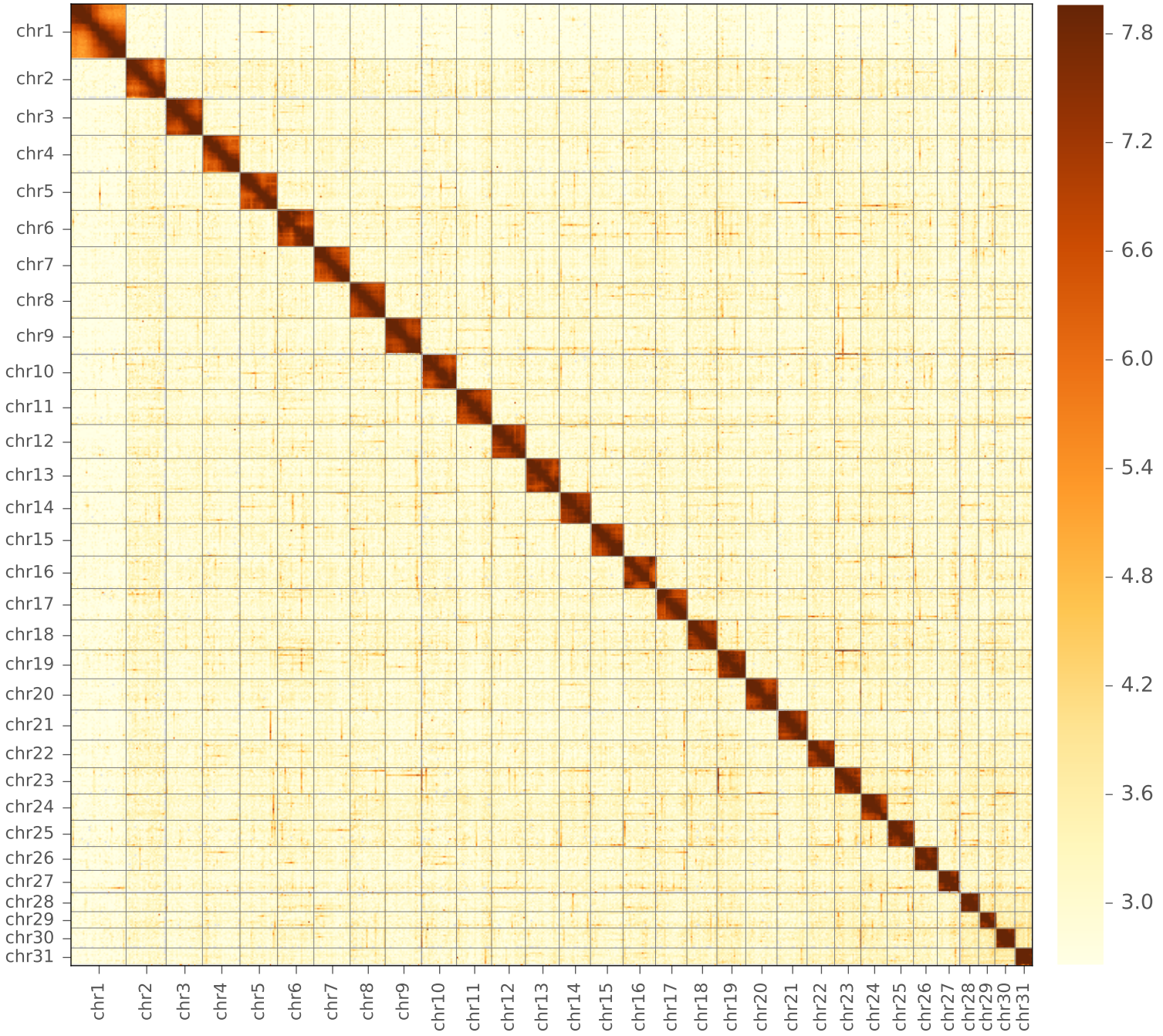
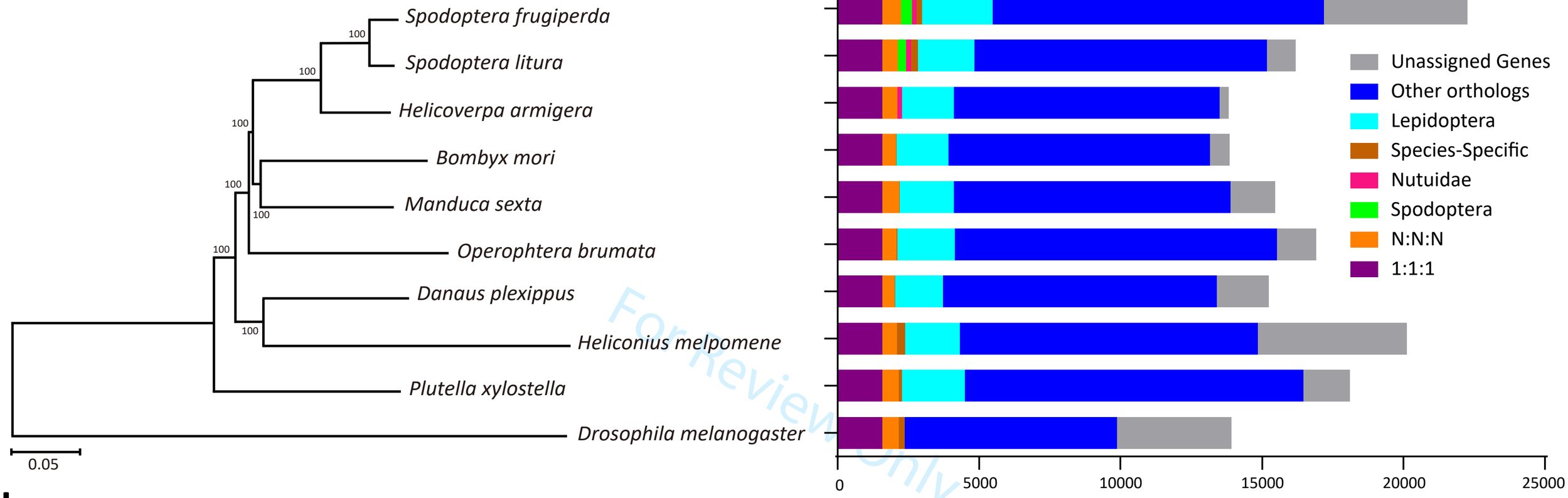


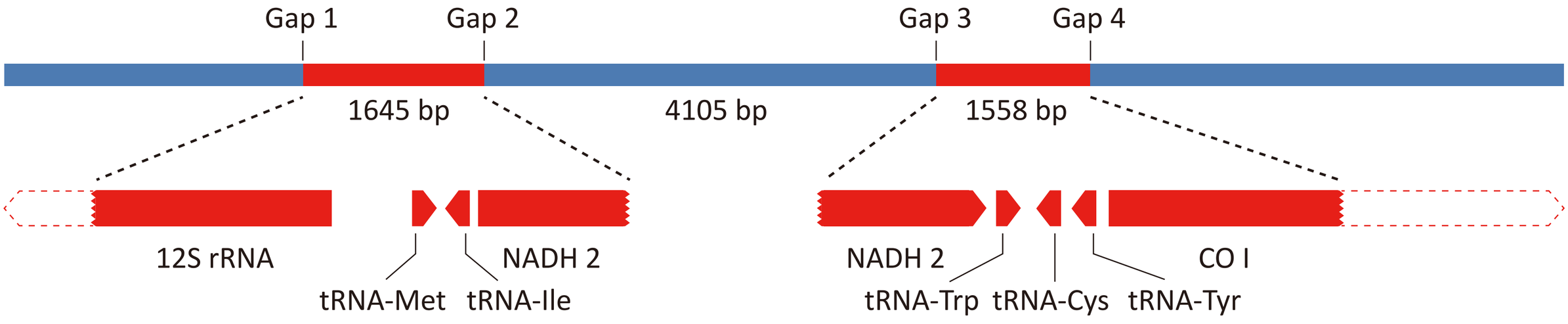
FIGURE 6 The LC_{50}s of two Chinese fall armyworm populations to different kinds of insecticides. cdcc and cdyc represent two inbred strains collected from Yunnan Province in China. Statistical significance of the difference was assessed by whether the 95% FL have overlap (*P <0.05).

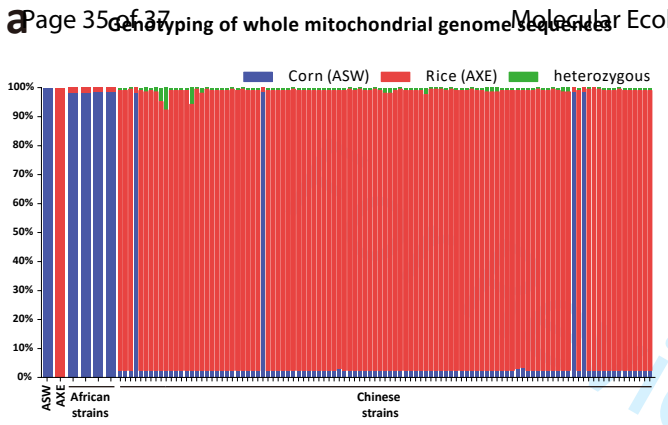

chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
chr23
chr24
chr25
chr26
chr27
chr28
chr29
chr30
chr31
7.8
7.2
6.6
6.0
5.4
4.8
4.2
3.6
3.0

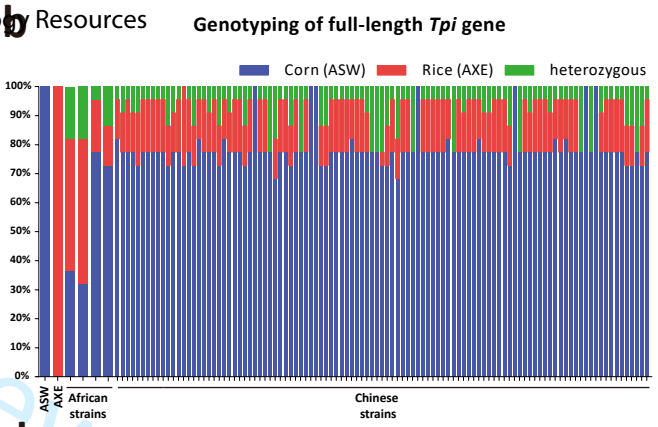
a

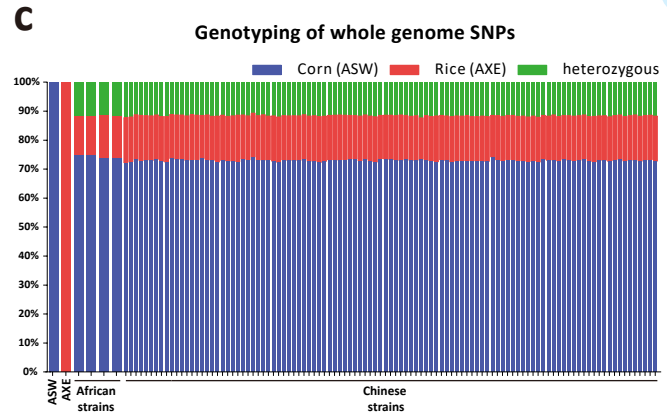


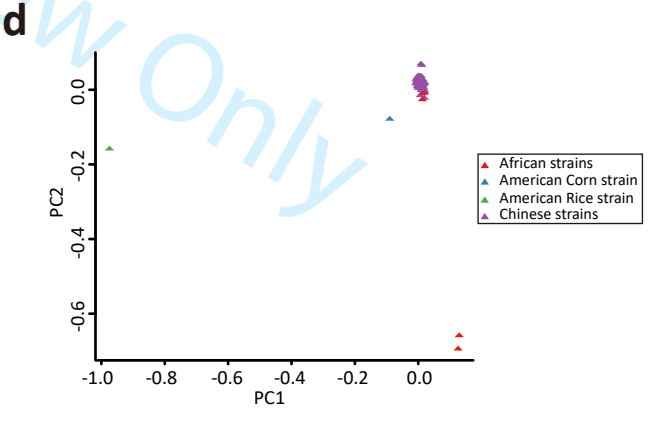
b

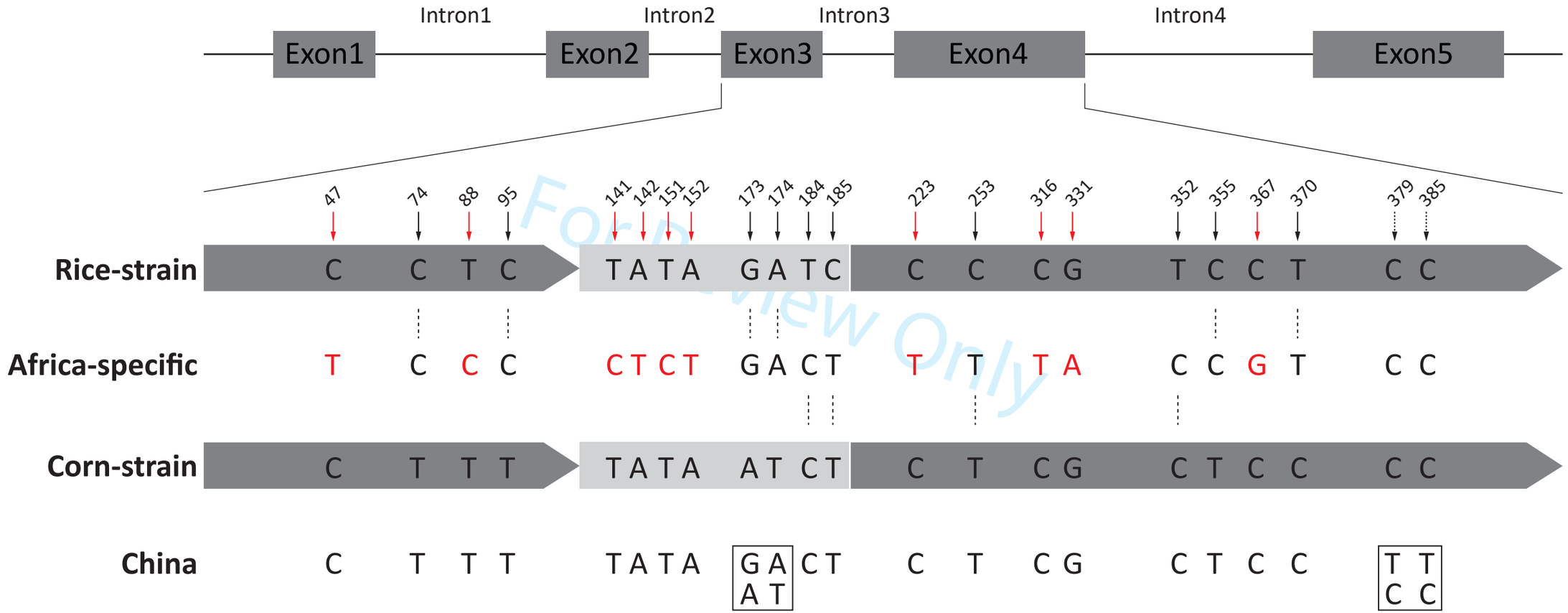








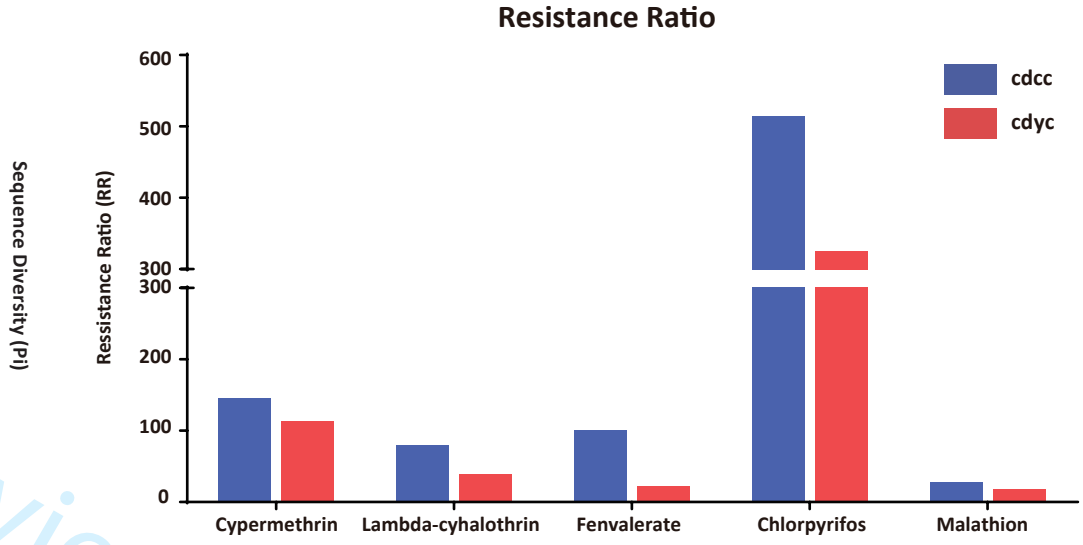




a

	AA201	AA227	AA290	
AChE-Susceptible:	A	G	F	
AChE-Resistance:	S	A	V	
Resequenced Samples 1:	A A	G G	F F	30.6%
2:	A A	G G	F V	39.6%
3:	A S	G G	F V	18.1%
4:	A A	G G	V V	11.7%

b



c