Global FAW population genomic signature supports complex introduction events across the Old World

Tay WT[1], Rane R[1], Padovan A[1], Walsh T[1], Elfekih S[2], Downes S[3], Nam K[4], d'Alençon E[4], Zhang J[5], Wu Y[5], Nègre N[4], Kunz D[6], Kriticos DJ[1], Czepak C[7], Otim M[8], Gordon KHJ[1].

1. CSIRO Black Mountain Laboratories, Clunies Ross Street, ACT 2602, Australia
2. CSIRO Australian Centre for Disease Preparedness, Geelong, Vic, Australia
3. CSIRO FD McMaster Laboratories, New England Highway, Armidale NSW2350, Australia
4. DGIMI, Université Montpellier, INRAE, Montpellier, France
5. College of Plant Protection Nanjing Agricultural University, Nanjing, China
6. Gordon Institute, University of Cambridge, Cambridge CB2 1QN, UK
7. Universidade Federal de Goiás, Escola de Agronomia, Goiânia, GO, Brazil
8. National Crops Resources Research Institute, Namulonge, Kampala, Uganda

Abstract

Accurate genomic knowledge can elucidate the global spread patterns of invasive pests. The high-profile invasive agricultural pest *Spodoptera frugiperda* (fall armyworm; FAW) is a case in point. Native to the Americas, the FAW was first reported in West Africa in 2016 and has rapidly spread to over 64 countries across the Old World, resulting in significant economic losses. The chronological order of reported detections has led to the hypothesis that the FAW moved eastwards across Africa and then Asia, however genomic evidence remains lacking to test this hypothesis and to identify the potential origin of invasive populations. Using a whole genome sequencing approach, we explored the population genomic signatures of FAW populations from the Americas and the Old World. Analyses of complete mitochondrial DNA genomes identified 12 maternal lineages across the invasive range, while genomic signatures from 870 high-quality nuclear genome-wide single nucleotide polymorphic (SNP) DNA markers identified five distinct New World populations that broadly reflected their native geographical ranges and absence of host-plant preference status. Unique FAW populations in the Old World were also identified that suggested multiple introductions underpinned their rapid global spread. We identified in Asian FAW individuals, genomes lacking evidence of admixture; while analysis of identified complex substructure revealed significant directional geneflow from Asia into East Africa, in contrast to a simple east-to-west spread. Our study highlights the need for population genomics approaches in analysing complex pest invasions, and the importance of international partnership to address global biosecurity challenges presented by emerging high priority insect pests.

Correspondence

Email: weetek.tay@csiro.au **Tel:** +61-2-6246 4286

Running title: FAW global spread by multiple introductions

Key words: Fall armyworm, global agricultural biosecurity, invasion biology, draft mitochondrial DNA genomes, single nucleotide polymorphic markers

Introduction

Global agriculture is increasingly affected by the spread of invasive pests and diseases, which is often assisted by global trade. A well-documented global invasion by an insect pest is that of the hemipteran sap-sucking *Bemisia tabaci* cryptic MEAM1 and MED species [1] that spread from the species' endemic population origins in the Middle East-Asia Minor and the Mediterranean region, respectively, to at least 54 countries [2]. Global agricultural trade has also been linked with the introductions of the highly polyphagous lepidopteran pest *Helicoverpa armigera* from the Old World [3] to South America [4,5] and the Caribbean islands [3,6-8]. This noctuid moth's wide host range, flight ability [9] and ability to develop resistance to insecticides [10] has enabled establishment across the New World with significant economic [11,12], and ecological consequences. Importantly, the availability of samples from early stages in the invasion has made it possible to distinguish demographic from adaptive evolutionary processes, as well as to study hybridization upon second contact with a sister species [13-16].

Long a significant pest of agriculture in its native New World range, the noctuid fall armyworm (FAW) *Spodoptera frugiperda* was first reported in West Africa (Nigeria and São Tomé and Príncipe) in early 2016 [17], followed by confirmation across central (Congo [18]; Togo [19], Southern [20] and Eastern [21]) sub-Saharan Africa between early 2017 to 2018 [22]. The Middle East (Yemen [23]) followed by India [24,25] and surrounding nations (e.g., Bangladesh, Sri Lanka). Myanmar [26] and Thailand [27] also reported detection of the pest between August and December 2018, followed by Southern China (Yunnan Province) in early January 2019 [28-30]. Detections of FAW since January 2019 have gathered speed: south-ward to Malaysia (March 2019) and Indonesia (Sumatra, April 2019; Java, July 2019; Kalimantan July 2019); Hong Kong (April 2019), Taiwan (May/June 2019); Laos and Vietnam (April 2019 [31]), the Philippines (June 2019 [32]; [33]), South Korea (June 2019), and Japan (June 2019) [34]. Within China, the FAW has been reported in a northward expansion pattern from Yunnan to 18 provinces by July 2019 [35-37]. As of August 2019, a total of 64 African and Asian nations have reported FAW [38]. In January 2020, FAW was trapped in Australia's special biosecurity zone in the Torres Strait islands of Saibai and Erub, and confirmed on 3 February 2020, and on mainland Australia in Bamaga on 18 February 2020 [39,40].

This chronologically ordered eastward spread of detections led to a widely adopted assumption [41] that the FAW was actually spreading west-to-east across and then from Africa. Based on the detection timeline, predictive simulations that assumed human-assisted spread, in particular agricultural trade, have modelled this very vagile pest's movement from the east coast of America/the Greater Antilles to West Africa (e.g., Togo [19]) between Central and Southern America and Africa, and between Africa and Asia (e.g., India, China; South East Asia [42]). The human-assisted spread model [42] was also used to warn China and South East Asian nations of imminent impact by FAW following confirmation of the pest in India [43]. This model further forms the basis of international research efforts to track the movement, including using molecular tools to examine invasion biology (e.g., [19,44,45]), and simulations to model long distance dispersal (e.g., [42,46,47]). Indeed a meteorological data-based simulation study (i.e., wind currents, monsoon wind patterns) concluded the Yunnan FAW populations originated from Myanmar, consistent with FAW being officially reported earlier in Myanmar (December 2018 [48,49]) than in China (January 2019 [35]). Other work has examined the impact and implications for global plant health and agricultural biosecurity (e.g., [28,50]), integrated pest management (IPM) and bioeconomics [51-53], and insecticide resistance [37,45,54].

Genetic studies on the spread of FAW have focussed on single genes on the mitochondrial genome, occasionally supported by a single partial nuclear gene marker. These markers have been widely used because, throughout much of the native range, FAW populations consist of two morphologically identical host races, the rice-preferred and corn-preferred *S. frugiperda* ('*Sfr*' and '*Sfc*', respectively), that have also been considered as potential sister species (e.g., [21,55,56]). These two host races are supported by phylogenetic analyses based on nuclear and mitochondrial DNA genomes [44], and partial mitochondrial DNA genes (e.g., [17,21,25,29,55]). The distribution of these *Sfr* and *Sfc* populations in their New World range has only recently been investigated based on partial mitochondrial and nuclear genes [57], while at the whole genome level they are less well-understood. Genotypes from both host races/sister species are present in the invasive populations (e.g., [19,58-60]). Since 2010 (e.g., [61,62]) and especially in recent times during the FAW range expansion (e.g., [29,45,59,63]), the partial *Triose Phosphate Isomerase* (*Tpi*) gene on the Z-chromosome has been adopted to aid in the clarification of the *Sfc* and *Sfr* host race status. The *Tpi* marker relies on the presence of a single nucleotide polymorphic (SNP) marker at position 183 (e.g., [62,63]) to differentiate between corn- or rice-preferred FAW. Similarly, inconclusive host preferences based on the mtCOI gene marker also detected both *Sfc* and *Sfr* on corn-host plants (e.g., [21]). Contrary to the introduction patterns of the related and invasive noctuid *H. armigera* in South America [4] which showed significant genetic diversity (e.g., [3,5,15]) similar to that reported for global populations of *H. armigera* [64-

[66], the current global partial mtCOI signatures of both *Sfc* and *Sfr* have been consistent with a single introduction, which, when considered together with the *Tpi* locus, was suggested to likely have a Florida/Greater Antilles source population [19].

What is missing from current research into the spread of FAW is analysis of broader genomic evidence. Genome-wide single nucleotide polymorphic (SNP) markers aligned to well-annotated genomes can provide powerful genomic evidence for understanding introduction pathways (e.g., [1]) and eliminate candidate populations [14] as well as elucidate hybrid signatures [13]. The genomes of both *Sfr* and *Sfc* have been sequenced and annotated [44], allowing higher resolution analysis of genetic structure, migration patterns and sub-species status based on a high number of genome-wide SNPs to enable identification of the potential New World origins, and the species and admixture status of the invasive *Sfc* and *Sfr* populations.

In this study, we provide an assessment of global FAW movement history based on genomic data that incorporates populations from both Northern, Central, and Southern Americas, and the Caribbean (i.e., representing the original population range), Western and Eastern Africa, and Western and Eastern Asia, representing the pest's Old World expansion. Here we reveal a multi-locus invasion that is likely independent of the reported detection patterns and their timelines, and provide genomic-based evidence to support multiple introductions of the FAW into the Old World, with significant movements of FAW detected between Asia and Africa. We also re-evaluated the pest's global spread directionality to highlight implications in future management of FAW, and the need for on-going global agricultural biosecurity research and cooperation to improve preparedness for emerging invasive agricultural pest issues.

Material and Methods

Spodoptera frugiperda populations sampled and analysed in this study were sourced from Florida (n=24) [67], Mississippi (n=18) [44], Puerto Rico (n=15) [68], Peru (n=16), Brazil (n=12; IBAMA Permit number: 18BR028445/DF), Mexico (n=10), Guadeloupe (n=4), French Guiana (n=3), Benin (n=4), India (n=12) [69]; Tanzania (n=1), Uganda (n=15), Malawi (n=16), and three populations from Yunnan Province, China (CC=19; CY=12; CX=15) [54], and one individual (CH06) from Australia's pre-border interception program overseen by the Department of Agriculture, Water and the Environment (DAWE), also from Yunnan, China (Suppl. Table 1). The initial differentiation of these individuals as 'corn-preferred' or 'rice-preferred' was based on the partial mtDNA COI gene region (see [55]) and a polymorphism within the Triose Phosphate Isomerase (*Tpi)* gene (see [62]).

Extraction of total genomic DNA was carried out at the CSIRO Black Mountain Laboratories site in Canberra Australia for the Brazil, Peru, Tanzania, Malawi and Uganda populations, as well as the pre-border intercepted FAW sample from China, using the Qiagen Blood and Tissue DNA extraction kit following instructions as provided, with genomic DNA eluted in 200µL EB. Total genomic DNA for the other three Chinese populations were extracted at Nanjing Agricultural University as detailed in Guan et al. [54]. Total genomic DNA from Mississippi, Florida, Puerto Rico, Guadeloupe, Mexico, and French Guiana, and Indian populations was carried out at INRAE DGIMI (Univ. Montpellier, INRAE, France) as described in Yainna et al. [69].

Genomic libraries prepared by CSIRO were constructed using an Illumina Nextera Flex DNA Library Prep Kit following manufacturer's instructions and sequenced by Illumina NovaSeq6000 S4 300 sequencing system at the Australian Genome Research Facility (AGRF). Sequencing efforts were shared between three research institutions: 61 samples were prepared at CSIRO (populations from Brazil, Peru, Uganda, Tanzania, and Malawi), 46 samples were prepared by NJAU for populations from China Yunnan Province (CC, CY and CX counties), and 89 samples were prepared by DGIMI, France (populations from Florida, Mississippi, Puerto Rico, Guadeloupe, French Guiana, Mexico, Benin and India). The Peru FAW samples and the single FAW sample CH06 from Yunnan China were intercepted at Australia's pre-border inspections of imported agricultural and horticultural commodities by the Department of Agriculture, Water and the Environment (DAWE) on fresh vegetables and cut flowers, respectively. The FAW CH06 was sequenced using the Illumina MiSeq high throughput sequencing (HTS) platform following the methods as described by Tay et al. [70].

Mitochondrial genomes assembly and haplotypes characterisation

The mitochondrial DNA genome for all samples were assembled using Geneious 11.1.5 based on strategies previously used for assembly of *Helicoverpa* species as outlined in Walsh et al. [71]. Assembled mitogenomes were annotated using MITOS [72] selecting invertebrate mitochondrial genetic code. All annotated protein coding

genes/coding sequences (PCGs/CDS) were re-annotated visually to identify putative stop codons and to align start codon positions. Four regions of low complexity (corresponding to BC55 nt6065-6092; nt9544-9580; nt12807-12838; nt15047-15276) were trimmed due to alignment difficulties and low genome assembly confidence associated with simple repeat units, resulting in all samples having final mitochondrial DNA genome length of 15,059 bp. We identified unique mitogenome haplotypes using the DNAcollapser in FaBox (1.5) <https://users-birc.au.dk/~palle/php/fabox/dnacollapser.php> [73] after alignment using MAFFT Alignment v7.450 [74,75] within Geneious 11.1.5 and selecting the Auto option for Algorithm and 200PAM / K=2 for Scoring matrix, Gap open penalty of 1.53, and offset value of 0.123.

SNPs selection

The reference genomes [44] are from native populations, *Sf*C from Guadeloupe and *Sf*R from Florida (see also Yainna et al. [69], and Nam et al. [68] for high quality assemblies of native population genomes for *Sf*R, and Nam et al. [76] for high quality genome assemblies of native *Sf*c). The reference genomes published recently by three Chinese groups may differ since they are from invasive populations, from Yunnan in China [77], from the Zhejiang province of China [78], and from 2017 sampled FAW individuals collected from maize fields in Lusaka, Zambia [45]. In this study, we used the original assembled genome of Gouin et al. [44] for our raw data processing. Genomic raw data was cleaned and trimmed using Trimmomatic and aligned to the *S. frugiperda* (rice v1) genome [44] using BWA-MEM [79]. Variants were predicted using BBMap [80] and indel's normalised using BCFtools [81] to obtain a whole genome SNP panel. Variants were filtered to remove SNP's with minimum allele frequency of 0.01 and linkage disequilibrium (LD) pruned to obtain 870 unlinked SNP's across all individuals.

Phylogeny analyses

Phylogenies based on trimmed partial mitochondrial DNA genomes of 15,059 bp and from genome-wide SPSs were individually inferred using IQ-Tree <http://iqtree.cibiv.univie.ac.at> [82]. For the nuclear SNPs, the panel of 870 SNPs from each individual in fasta format was up-loaded to the IQ-Tree web server and selecting the automatic substitution model option. For the mitochondrial DNA genome Maximum Likelihood (ML) phylogeny was inferred with edge-linked partition for the 13 protein coding genes and excluding all four regions of low complexity. We used the Ultrafast bootstrap analysis [83] with 1,000 bootstrap alignments to assess branch support for both mitochondrial DNA genome and nuclear SNPs phylogenies. Output consensus tree files in Newick format were visualised and manipulated using Dendroscope version 3.5.7 [84].

Genetic diversity and neutrality tests

Observed (H_{obs}) and expected (H_{exp}) heterozygosity were calculated for each population using the populations program in Stacks [85] and the Adegenet package in R [86,87]. The number of loci departing significantly from Hardy-Weinberg equilibrium (HWE) in the global population and individual populations was assessed using PLINK 2.0 [88] and VCFtools [89]. To test for neutrality, Tajima's D [90] and Fu and Li's D* [91] were calculated for each population using the PopGenome package in R [92]. Nucleotide diversity (π) and Wright's inbreeding coefficient, F_{IS} [93], were calculated using the populations program in Stacks. Pairwise comparisons of weighted F_{ST} values between populations were calculated using Genepop (v4.7.5 [94]) and differentiation between populations tested for significance using the exact G test.

Population structure and migration

Principal component analysis (PCA) was performed using PLINK v1.9 [95]. Admixture was estimated using Admixture v1.3.0 [96]. For plotting of networks, the R package NetView [97,98] was used. The network drawn using the plotAdmixture function in this package is based on a Maximum Likelihood (ML) distance matrix calculated from the IQ-Tree phylogeny shown in Fig. 3, using the R package 'ape' [99].

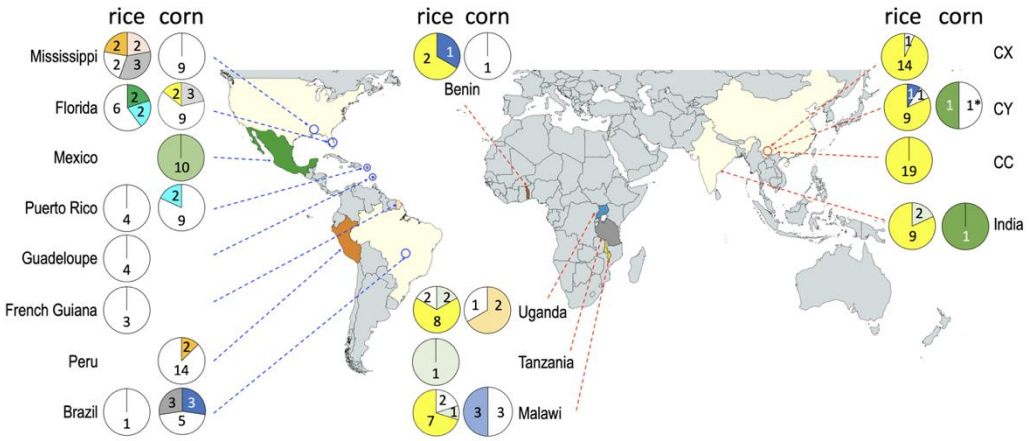
To estimate directional gene flow between the populations, as well as the relative magnitudes of these flows, the function divMigrate in the R package diveRsity [100] online version was used <https://popgen.shinyapps.io/divMigrate-online/> [101]. Gene flows between all sites were calculated and then normalized to obtain relative migration rates (between 0 and 1). The program divMigrate searches for gene flow directionality between each pair of populations by identifying significant asymmetry using a hypothetically

defined pool of migrants for each pair. This pool is then compared to pairs in the overall population to calculate directional genetic differentiation and relative migration. To evaluate the significance of asymmetric migration, 1,000 bootstraps were performed. Resulting migration matrices were then plotted using Gephi <https://gephi.org/> to generate network graphs. These show directional gene flows between populations (located at the nodes), with the thickness of the lines showing relative strength of gene flow.

Results

A total of 197 FAW individuals were sequenced, 102 from the native New World range and 95 from the invasive Old World range (Fig. 1). From the pest's native range, we detected 25 'rice' mitochondrial genome (i.e., mitogenome) haplotypes, and 51 'corn' mitogenome haplotypes. All FAW from Mexico and Peru had the 'corn' mitogenome while FAW from Guadeloupe and French Guiana were all 'rice' mitogenomes. Of the FAW from the invasive range nine 'corn' and 'rice' mitogenome haplotypes were identified; one of the 'corn' mitogenome haplotype (represented by green colour, Fig. 1) was shared between CY and Indian individuals. No African corn mitogenome haplotypes were shared with Asian FAW populations. In contrast, 83% (i.e., 68/82) of African and Asian FAW with 'rice' mitogenomes shared a common haplotype (represented by the yellow colour, Fig. 1). FAW individuals from China and Benin shared a rare rice mitogenome haplotype (blue colour haplotype), and individuals from Uganda, Tanzania, Malawi and India shared one unique (i.e., light green colour) haplotype (Fig. 1). In general, the high diversity of haplotypes in both 'rice' and 'corn' in the native range and the lack of diversity in the invasive range is consistent with patterns observed in invasive populations that have a relatively small number of founders. GenBank accession numbers for full mitochondrial genomes from all individuals are listed in Suppl. Table 1.

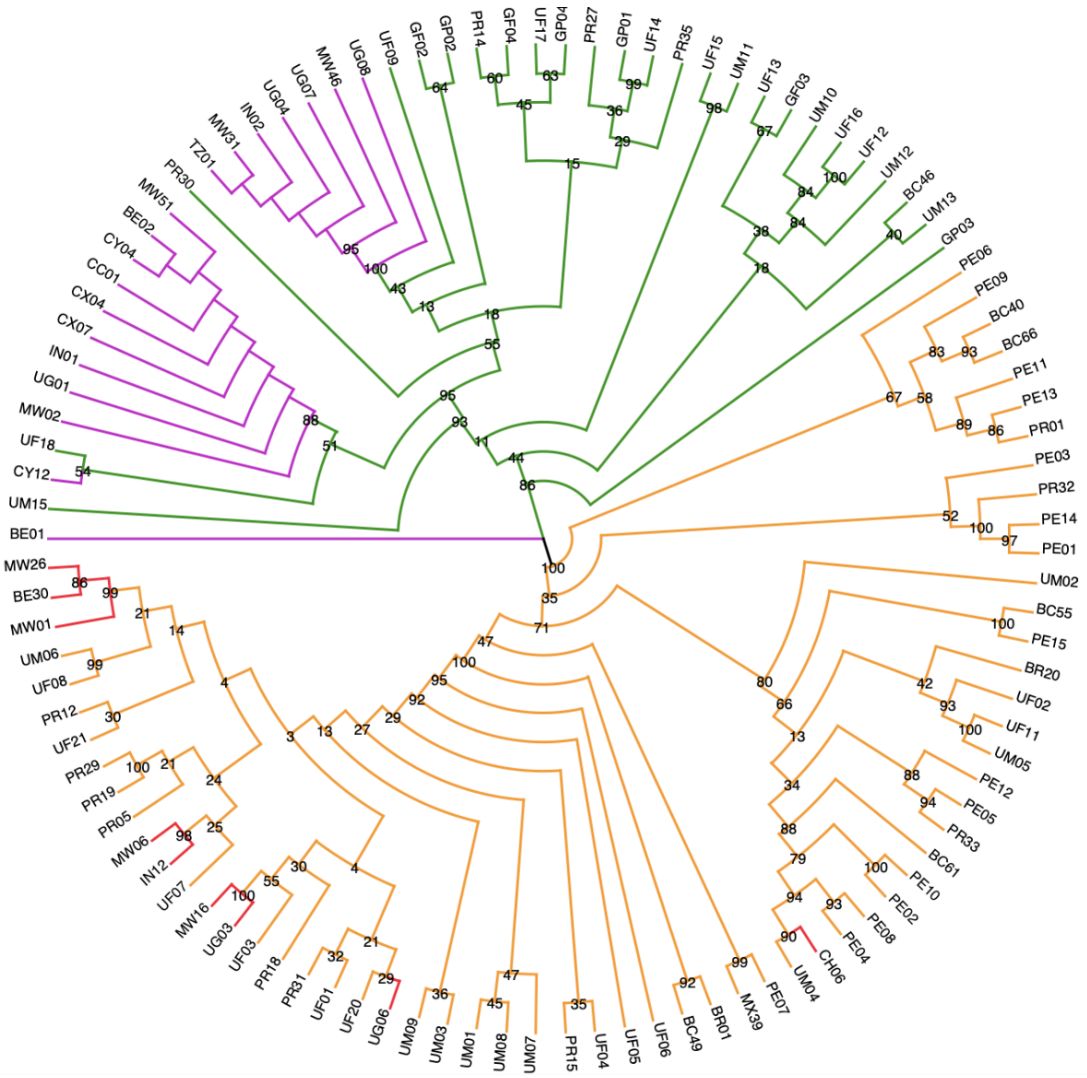
Fig. 1: New and Old Worlds' FAW populations and proportions of mitochondrial DNA haplotypes based on 15,059 bp of the mitochondrial DNA genomes and excluding four regions of low complexity. For the New World 'rice' FAW, a total of 20 unique mitogenome haplotypes (represented by white colour proportion of each pie chart), and 11 non-unique mitogenome haplotypes were detected (i.e., a total of 25 mitochondrial haplotypes in rice FAW in the New World). For the 'corn' mitogenomes, 46 unique haplotypes were detected from the native range, while 25 corn FAW individuals shared a total of seven haplotypes (i.e., a total of 46+7=53 mitochondrial haplotypes). In the invasive range, six unique 'rice' mitogenomes (i.e., white portion of the pie charts, representing two individuals from Uganda, two individuals from Malawi, and two individuals from China (CY, n=1; CX, n=1) and three shared mitogenomes (i.e., dark blue, yellow, pale green) were detected from 76 individuals from Africa (n=22), India (n=11) and China (n=43). For the 'corn' FAW from the invasive range, six unique mitogenome haplotypes (i.e., white portions of pie charts) and three non-unique mitogenome haplotypes (pale orange, pale blue and dark green) were detected, although only one individual each from China and India shared a common mitogenome (represented by dark green). With the exception of white colour representing unique mitogenomes, colour schemes are otherwise independent between 'corn' and 'rice' mitogenome haplotypes. China FAW populations from Yunnan Province of Cangyuan (CC), Yuanjiang (CY), and Xinping (CX) are indicated. One pre-border FAW intercepted on December 2016 from cut flowers that originated from Yunnan China (CH06) with a unique corn mitogenome is indicated with '*' (placed together with the CY corn pie-chart). Numbers within pie-charts indicate individuals for each mitogenome haplotype.



Mitochondrial DNA genome phylogeny

The trimmed (15,059bp) mitochondrial DNA genome phylogeny of all individuals in our study identified two distinct clades that corresponded to the 'rice-preferred' and 'corn-preferred' clusters (Fig. 2). Based on the near complete mitogenome phylogeny, a minimum of four and five introduction events were likely associated with the 'rice' and 'corn' maternal lineages, respectively (Fig. 2). Except for the 'corn' specimen (CH06) from Yunnan that clustered strongly with an individual from Mississippi (UM04), all 'corn' individuals from the invasive range (i.e., MW26, BE30, MW01, MW06. IN12, MW16, UG03, UG06) clustered weakly with individuals from Florida. Similarly, apart from the Benin individual (i.e., BE01), all remaining 'rice' FAW from the invasive range also clustered weakly with individuals from Florida. Therefore, the likely origins of the invasive 'corn' and 'rice' FAW in the Old World is inconclusive based on the near complete draft mitogenome phylogeny.

Fig. 2: FAW Maximum Likelihood phylogeny constructed using IQ-Tree based on 15,059 bp partial mitochondrial genome with edge-linked partition for the 13 protein coding genes and excluding four regions of low complexity. Node support estimated from 1,000 bootstrap replications. 'Rice' clade is indicated by branches in green (native range) and purple (invasive range), and 'corn' clade is indicated by branches in orange (native) and red (invasive range). Unique haplotypes from all populations are included. Country codes are: UF (USA-Florida), UM (USA-Mississippi), PR (Puerto Rico), GP (Guadeloupe), GF (French Guiana), PE (Peru), MX (Mexico), BC (Brazil-CC), BR (Brzil-rCC), BE (Benin), UG (Uganda), TZ (Tanzania), MW (Malawi), IN (India), and four populations from China Yunnan Province (Australia pre-border interception (CH06); Cangyuan (CC), Yuanjing (CY), and Xinping (CX)).



Nuclear SNP Phylogeny

The ML phylogeny based on 870 unlinked and neutral SNPs revealed four distinct clades (clades I, II, III, IV; see Fig. 3) across the sampled populations. Native and invasive individuals were a component of each clade which enabled a side-by-side comparison of population structure. Members within each clade were grouped with high (90-96%) bootstrap branch node support values. Clade I included the majority of the invasive FAW individuals from China (CX, CY, CC populations), India (IN), Uganda (UG), and Benin (BE) as well as individuals from Brazil. Overall, subclades within Clade I indicated unique genomic signatures between the CC and CY/CX populations. Indian and African populations (i.e., Uganda, Benin) were scattered among the CC and CY/CX populations. This interspersed clustering of subclades from Chinese, African and Indian populations suggests a complex FAW spread across the Old World, with some of the China CY individuals potentially sharing a New World origin similar to the Brazil rCC (i.e., 'BR' code, Fig. 3 Clade I) individuals.

Clade II, which is phylogenetically most closely related to Clade I, is dominated by individuals from Mississippi. Within this clade, individuals from China (i.e., CX), Uganda, Benin and India are also present, indicative of likely separate introductions of FAW from population(s) with genetic similarity to the Mississippi population into the Old World. Clade III is represented by a separate Brazilian (i.e., 'BC') FAW population and the Peru FAW individuals. Invasive populations clustered within clade III were the Malawi FAW population, a single Tanzania and three Ugandan individuals, suggesting that these African FAW shared a similar origin that is different from other African (e.g., Benin, rest of Uganda) and Asian populations. The Ugandan population in particular appears genetically highly heterogeneous, indicating it too have mixed introduction backgrounds.

Clade IV is dominated by the Florida population and other Caribbean islands/Greater Antilles (e.g., Puerto Rico)/Lesser Antilles (e.g., Guadeloupe)/ Central American (e.g., Mexico), and parts of the northern region of South America (e.g., French Guiana) FAW populations. Clade IV contained a single invasive Chinese FAW (i.e., CH06; intercepted at Australia's pre-border inspection program). Taken as a whole, the nuclear SNP phylogeny provides clear evidence for multiple introductions of FAW to the Old World (African and Asian continents), while identifying populations associated with the Mississippi and the Brazilian 'BR' populations as likely sources of invasive populations into the Old World. The source population for Malawi's FAW was likely population(s) from South America, currently represented by Peru/Brazil (BC) populations. Based on interception data, with the exception of a single unique FAW, Florida and the Greater Antilles do not appear to be likely sources for the current invasive populations in the Old World.

Our nuclear SNP phylogeny therefore clearly showed that the native range FAW populations could be classified based on their geographic origins. The unexpected direct phylogenetic relationship between the US Mississippi and Brazil 'BR' population, suggested potential movements of populations within North America (i.e., Mississippi is not the wintering ground for FAW and represents the melting-pot for migrating individuals from Texas and Florida that are only present in the summer seasonal migration [102]) and between North and South America. Finally, a significant overall finding was that our panel of neutral SNPs selected from whole genome sequencing did not separate individuals based on 'corn' or 'rice' mitochondrial DNA genome signatures, nor did they support the host strain characterisation based on the *Tpi* partial gene marker (see below).

Genetic diversity and neutrality tests

Basic population diversity statistics for each population are listed in Table 1. Nucleotide diversity (π) varied across a narrow range (0.287-0.329), for the limited number of variable and independent SNPs analysed, that included no invariant loci. No significant overall difference was observed between the native and invasive range populations. All populations showed higher average observed heterozygosity (H_{obs}) than the average expected heterozygosities (H_{exp}), both in the native and invasive ranges, with the highest H_{obs} seen in the Malawi population. Negative F_{IS} values for all populations were consistent with H_{obs} being higher than H_{exp}, and suggested systematic avoidance of consanguineous mating [103] within FAW subpopulations as a whole. The heterozygosity excess observed in all these populations is most likely indicative of the recent mixing of previously distinct populations. These are likely the result of multiple introductions into the invasive range, as already suggested by the nuclear SNP phylogeny (Fig. 3) and PCA (Fig. 4). The heterozygosity excess detected for the native range populations may similarly be due to high levels of migration among populations in the native range. Consistent with these observations, a number of the populations including most from the invasive range also contained significant numbers of loci not in Hardy-Weinberg equilibrium (HWE). This was especially the case for the two largest populations from China (i.e., CY and CX), Malawi and Uganda, as well as for several native range

populations; many of the populations studied therefore appear to result from recent mixing of previously separated populations. Approximately half the number of loci departed significantly from HWE in the global population (i.e., 437 of the total 870) and highlighted the complex global population structure.

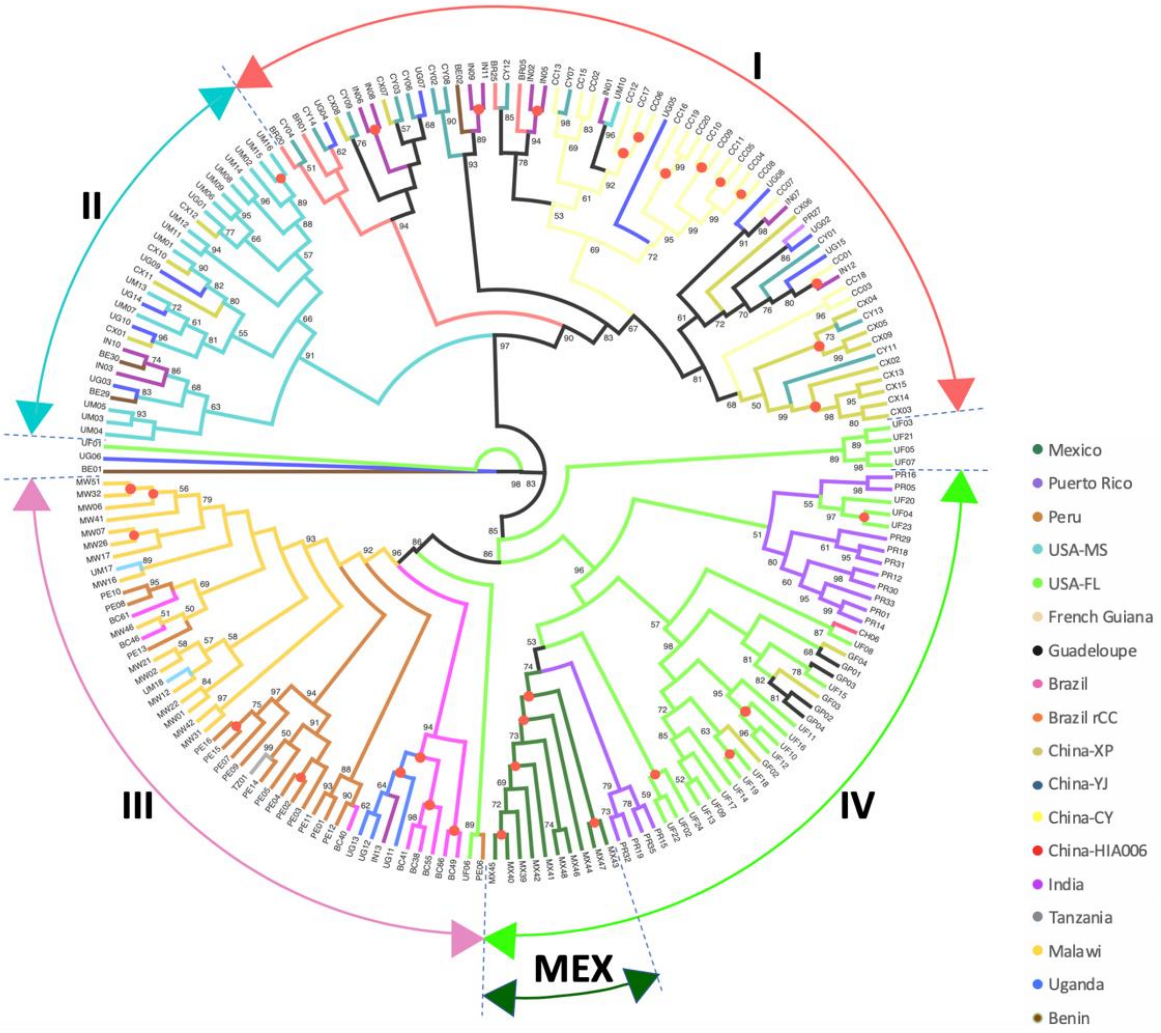
The statistical tests for neutrality were also highest for those populations with highest H_{obs} and lowest numbers of loci in HWE. For the Tajima's D estimates, all populations showed positive D values. Some populations showed lower Tajima's D values (below 0.4), suggesting that these populations were evolving at close to mutation-drift equilibrium. The high Tajima's D calculated for most of the populations (particularly values > 1.40) supports them having structure that likely reflects each comprising multiple recent introductions from different source populations. The other test for neutrality, Fu & Li's D*, gave consistent results with Tajima's D, further supporting the finding of genomic signatures in these populations due to recent introduction of different populations. These demographic factors are the likely reason for the elevated neutrality test results, since these were averages across the complete SNP set derived from the whole genome, rather than results for individual loci that might be under balancing selection [69]. Interestingly, this may also apply to populations in the native range. Another potential explanation for the high D and D* values, that of rapid population collapse, is unlikely since it would require much longer evolutionary time frames than in the current populations sampled.

Table 1: Population statistics for native and invasive range FAW populations.

Pop. Code	Pop.	No. Samples	Avg. H_{exp}	Avg. H_{obs}	HWE, P>0.001	Tajima's D	Fu & Li's D*	F_{IS}	Nt diversity (π)
BC	Brazil-CC	8	0.289	0.420	870	0.693	0.671	-0.241	0.309
BE	Benin	4	0.274	0.408	870	0.331	0.328	-0.179	0.313
BR	Brazil-rCC	4	0.263	0.396	870	0.130	0.130	-0.178	0.301
CC	China-CY	19	0.282	0.400	796	1.418	1.286	-0.262	0.289
CH	China-H	1							
CX	China-XP	15	0.293	0.416	837	1.249	1	-0.263	0.303
CY	China-YJ	12	0.284	0.405	870	1.013	1	-0.248	0.296
GF	French Guiana	3	0.247	0.375	870			-0.138	0.296
GP	Guadeloupe	4	0.245	0.359	870	0.293	0	-0.152	0.279
IN	India	12	0.289	0.403	870	1.137	1	-0.239	0.301
MW	Malawi	16	0.319	0.461	838	1.411	1.30	-0.303	0.329
MX	Mexico	10	0.265	0.403	870	0.717	0.610	-0.263	0.279
PE	Peru	16	0.319	0.456	848	1.554	1.293	-0.295	0.329
PR	Puerto Rico	15	0.288	0.404	845	1.241	1.188	-0.251	0.298
TZ	Tanzania	1							
UF	USA-FL	24	0.281	0.383	810	1.470	1.699	-0.242	0.287
UG	Uganda	15	0.305	0.428	843	1.795	1.298	-0.266	0.315
UM	USA-MS	18	0.320	0.453	820	1.717	1.452	-0.293	0.329

Note: The native range FAW populations are: USA-Florida (UF), USA-Mississippi (UM), Brazil-rCC (BR), Brazil-CC (BC), Puerto Rico (PR), Guadeloupe (GP), French Guiana (FG), Peru (PE), Mexico (MX) and the invasive range FAW populations are Benin (BE), Uganda (UG), Tanzania (TZ), Malawi (MW), India (IN), and China (CH, CC, CY, CX). See Suppl. Table 1 for sample and population details, and see Materials and Methods for details of how the statistics were calculated. Neutrality tests (Tajima's D and Fu & Li's D*) were only calculated for populations with at least 4 samples. Nucleotide diversity was calculated using Stacks only for the variant loci analysed and no window size specified.

Fig. 3: IQ-Tree with 1,000 bootstraps replications to estimate node support for *Spodoptera frugiperda* populations from Northern America (Mississippi, Florida), Caribbean (Puerto Rico, Guadeloupe, French Guiana), and South America (Peru, Brazil), as well as *S. frugiperda* populations representing the Old World invasive range from Western Africa (Benin), Eastern Africa (Uganda, Tanzania, Malawi), and Asia (India, China). A total of 870 independent SNPs (i.e., unlinked) from non-coding regions distributed across the genome with no missing data were used. Populations are represented by unique colour schemes as indicated. Three populations of *S. frugiperda* from China Yunnan Province are Cangyuan (CC), Yuanjiang (CY), and Xinping (CX), and two populations of *S. frugiperda* from Brazil are Brazil-CC (BC) and Brazil-rCC (BR). Branch nodes with 100% bootstrap support are indicated by red dots. Bootstrap values of <50% are not shown.



Population structure and migration

Multivariate Principal Component Analysis (PCA) of the 197 individuals in the native and invasive populations was also based on the 870 neutral and unlinked SNP loci and showed the individuals to largely cluster according to their populations, as previously observed in the phylogenetic analyses (above). The native FAW populations formed five clusters as shown in Fig. 4A. More detailed analysis of the native range samples showed those from Peru to clustered overall with the Brazil-CC population (code 'BC') while also overlapping with those from Florida (Fig. 4B), and the samples from Puerto Rico, Guadeloupe and French Guiana tended to cluster with Floridian population with 96% confidence (Fig. 4C). This panel also showed the invasive FAW population from Malawi clustering with Brazil-CC and Peru in Clade III, with 96% confidence. The Ugandan population was scattered across Clades I, II and III (Fig. 4D) while the Benin individuals fell within clades I and II

and that from Tanzania fell just outside of 96% confidence of Clade III. Indian FAW individuals showed similar clustering patterns to those of Ugandan individuals, being found in Clades I, II, and III (Fig. 4E). The Chinese FAW populations were predominantly clustered within Clade I, with a few XP individuals also found within Clade II (Fig. 4F). No individual from China was found in Clade III, while one individual (CH06) was clustered with Florida population (Clade IV) at 96% confidence. We did not identify any invasive population to cluster with the Mexican population.

Pairwise gene flow estimates (F_{ST}) between the populations varied significantly (Table 2). The Mexico and Brazil-rCC (BR) populations overall showed very limited gene flow with all other populations, while the Brazil population showed a low level of gene flow with both Peru and US Mississippi (UM) populations. There was a lack of population substructure especially between invasive range populations which suggests varying levels of gene flow. Significant population substructure was detected between Peru and invasive FAW populations from China-CY, China-XP and China-YJ, and India, while F_{ST} estimates indicated gene flow occurring between Peru and African populations (Benin, Tanzania, Uganda, and Malawi), suggesting some level of movements within African populations.

Admixture analysis

Analysis of populations using Admixture showed structure evident at K values from 3 to 5 (Fig. 5). At K=3, a total of six Chinese individuals from the CY and YJ populations appeared to be non-admixed (grey dots). Similarly, at K=4, three of these six FAW individuals remained non-admixed as indicated by grey dots. However, at K=5, the number of non-admixed individuals nearly doubled compared with K=3, with individuals from CC (indicated by grey dots) and CX (indicated by red dots) being most different overall, although there was also one individual from CX showing a non-admixed genome signature similar to those detected in the CC population, and one CY individual with a non-admix genome pattern similar to those detected in CX. No other FAW individuals from the invasive range otherwise showed non-admixed genomic signatures irrespective of the K-values of 3, 4 or 5. As expected based on PCA and nuclear SNP phylogeny findings, the Malawi FAW individuals share very similar admixture patterns as FAW individuals from Peru and Brazil-CC (i.e., 'BC') populations. This shared admixed profile between Malawi and Peru/BC populations is especially clear at K=5, which also enable clearer visualisation of the Tanzanian individual and selected Ugandan individuals (e.g., UG11, UG12, UG13) as also having similar admixture profiles as Malawi individuals.

Admixture analysis of native populations of FAW showed that majority of individuals have admixed genomic signatures. The exceptions being individuals from Florida (e.g., UF19, UF09, UF12, UF16), and Guadeloupe (GP02, GP04) at predominantly K=4 and K=5. Interestingly, these individuals with non-admixed genomic signatures (at either K=3, 4 and/or 5) also possessed the rice mitogenome haplotypes (Fig. 5). This observation is similar to that observed for the non-admixed Chinese individuals that have mitogenomes which also exhibited the rice haplotypes. Admixture analysis also revealed most Mexican individuals as having non-admixed genome patterns and with the corn mitogenome haplotypes (Suppl. Figs. 1A and 1B). As found in the phylogenetic analyses (above), comparison of the admixture patterns to mitogenomes and the *Tpi* locus of native and invasive FAW populations failed to identify any evidence to support the characterisation of FAW, and especially in the invasive range, as either 'corn-' or 'rice-' preferred *S. frugiperda*. The genome admixture signatures of FAW across its African and Asian invasive range again suggest a complex pattern for FAW introduction into the Old World. For example, given the highly admixed genomic patterns detected in African and Indian FAW individuals, it is unlikely that matings between these admixed populations would lead to individuals with non-admixed genomic signatures in China unless there was some very strong selection pressure acting across these selected CY, CC, and CX individuals' genome as a whole.

Table 2: Population pairwise gene flow estimates (F_{ST}) between native and invasive range FAW populations.

	BE	BC	BR	CC	CH	CX	CY	GF	GP	IN	MW	MX	PE	PR	TZ	UF	UM	UG
BE	N/A											***						
BC	0.027	N/A		***		***	***			***	^	***		***		***		***
BR	0.071	0.039	N/A	***		***	***			***	+	***		+		***		***
CC	0.012	0.05	0.09	N/A		***			***		***	***	***	***		***	***	***
CH	0.029	0.02	0.124	0.057	N/A													
CX	0.019	0.05	0.08	0.02	0.049	N/A			***		***	***	***	***		***	***	
CY	0.006	0.04	0.08	0.008	0.052	0.012	N/A		^		***	***	***	***		***	***	
GF	0.036	0.039	0.088	0.048	0.035	0.053	0.052	N/A				***						
GP	0.041	0.048	0.086	0.06	0.046	0.06	0.05	0.002	N/A		+	***	^				***	
IN	0.004	0.04	0.07	0.008	0.032	0.012	0.006	0.037	0.043	N/A	^	***	***	***		***	***	
MW	0.007	0.02	0.05	0.02	0.024	0.03	0.02	0.039	0.05	0.02	N/A	***		***		***	***	
MX	0.07	0.07	0.1	0.08	0.088	0.08	0.08	0.09	0.09	0.07	0.07	N/A	***	***		***	***	***
PE	0.02	0.011	0.042	0.05	0.018	0.05	0.04	0.037	0.04	0.04	0.014	0.06	N/A	***		***		***
PR	0.018	0.03	0.06	0.03	0.024	0.03	0.03	0.031	0.039	0.02	0.028	0.05	0.03	N/A		***	***	***
TZ	0.041	0.051	0.14	0.071	Nan	0.076	0.069	0.081	0.098	0.047	0.007	0.132	0.019	0.066	N/A			
UF	0.016	0.03	0.06	0.03	0.016	0.03	0.03	0.008	0.014	0.02	0.03	0.06	0.03	0.01	0.062	N/A	***	***
UM	0.034	0.018	0.042	0.05	0.026	0.05	0.05	0.041	0.05	0.04	0.02	0.06	0.01	0.03	0.021	0.03	N/A	***
UG	0.003	0.03	0.06	0.02	0.025	0.021	0.013	0.032	0.043	0.011	0.001	0.07	0.02	0.02	0.011	0.03	0.03	N/A

Note: The populations are denoted as in Table 1. The F_{ST} values are given in the lower left half of the table, and the p-values (*** $p << 0.001$; + $p \leq 0.01$; ^ $p \leq 0.05$) in the upper right. Both Tanzania and China-H06 populations consisted of one individual each and their pairwise F_{ST} was therefore not estimated.

Fig. 4: Principal Component Analyses of native and invasive FAW populations based on 870 neutral and unlinked SNP loci. **Panel A** shows the five clusters of native FAW populations (identified also from the genome-wide SNP phylogeny in Fig. 3). Circles indicate confidence as shown in Fig. 3. **Panel B:** Peru individuals clustered overall with Brazil-CC population (Clade III; pink colour) but also overlapped Florida population (Clade IV, light green colour). **Panel C:** Puerto Rico (purple colour), Guadeloupe (black colour) and French Guiana (wheat) overall clustered with Florida population with 96% confidence, while the invasive FAW population from Malawi (yellow colour) clustered in Clade III with Brazil-CC and Peru with 96% confidence. **Panel D:** PCA of Uganda population (blue colour) indicated the population were scattered across Clades I, II and III, Benin individuals (Saddlebrown colour) fell within clades I and II, while Tanzania (Azure 4 colour) fell just outside of 96% confidence of Clade III. **Panel E:** Indian FAW individuals showed similar clustering patterns as the Ugandan individuals, being found in Clades I, II, and III. **Panel F:** Chinese FAW populations were predominantly clustered within Clade I, with few CX individuals also found within Clade II. No individual from China was found in Clade III, while one individual originating from Australia's pre-border inspection program was clustered with Florida population (Clade IV) at 96% confidence. No invasive populations were clustered with the Mexican population. Colour codes for populations as provided in Fig. 3.

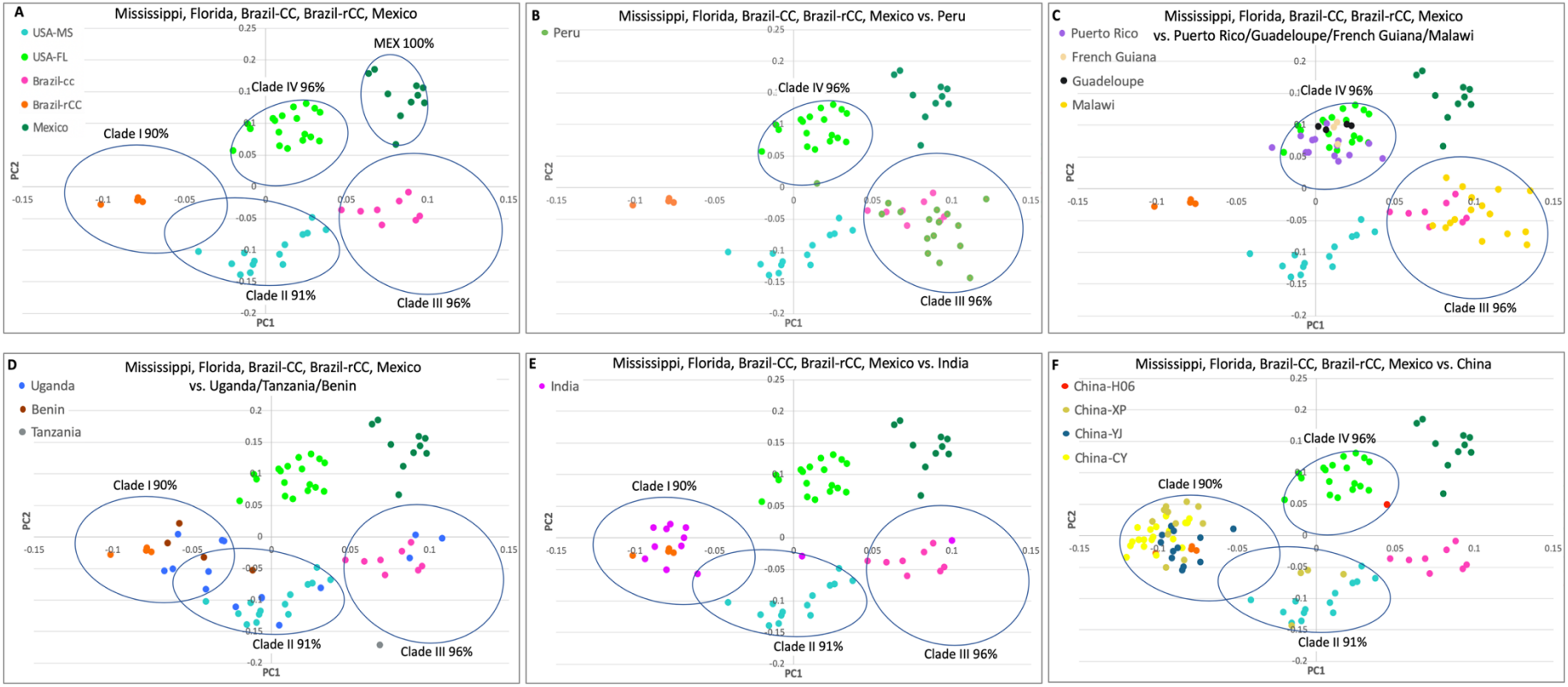
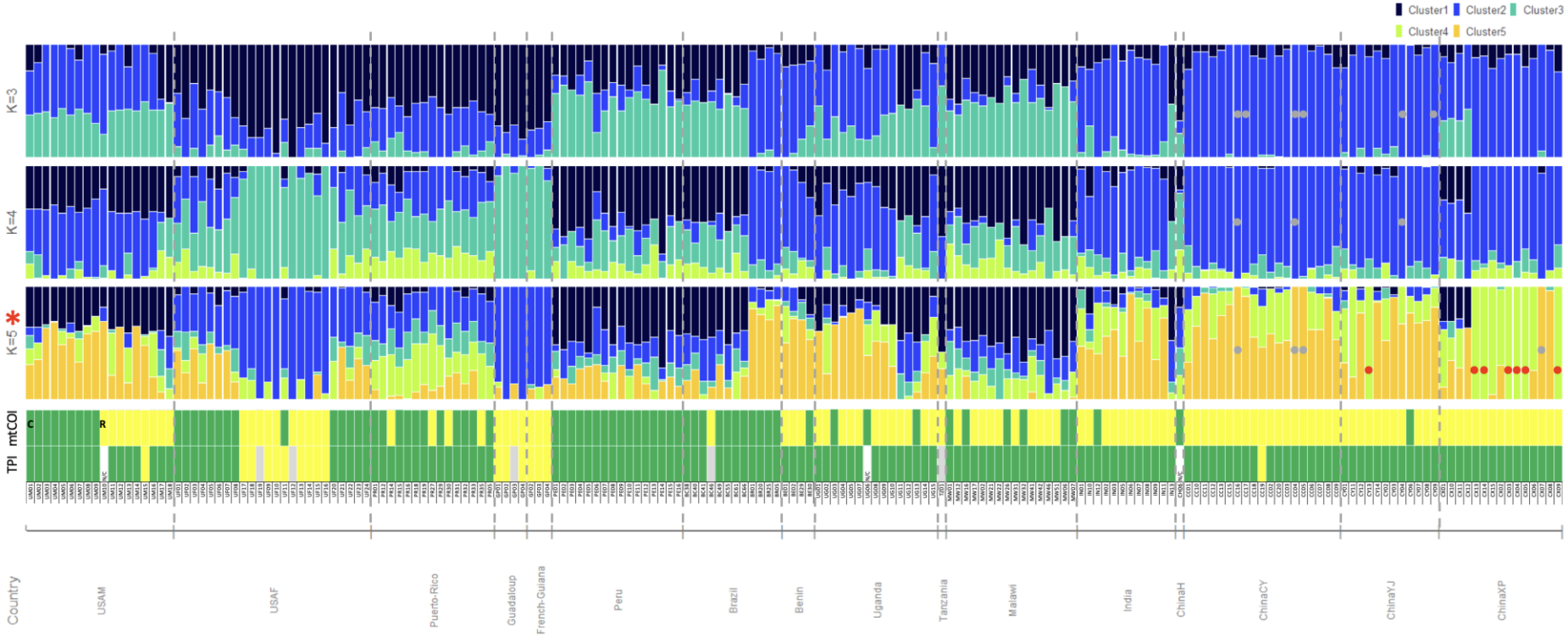


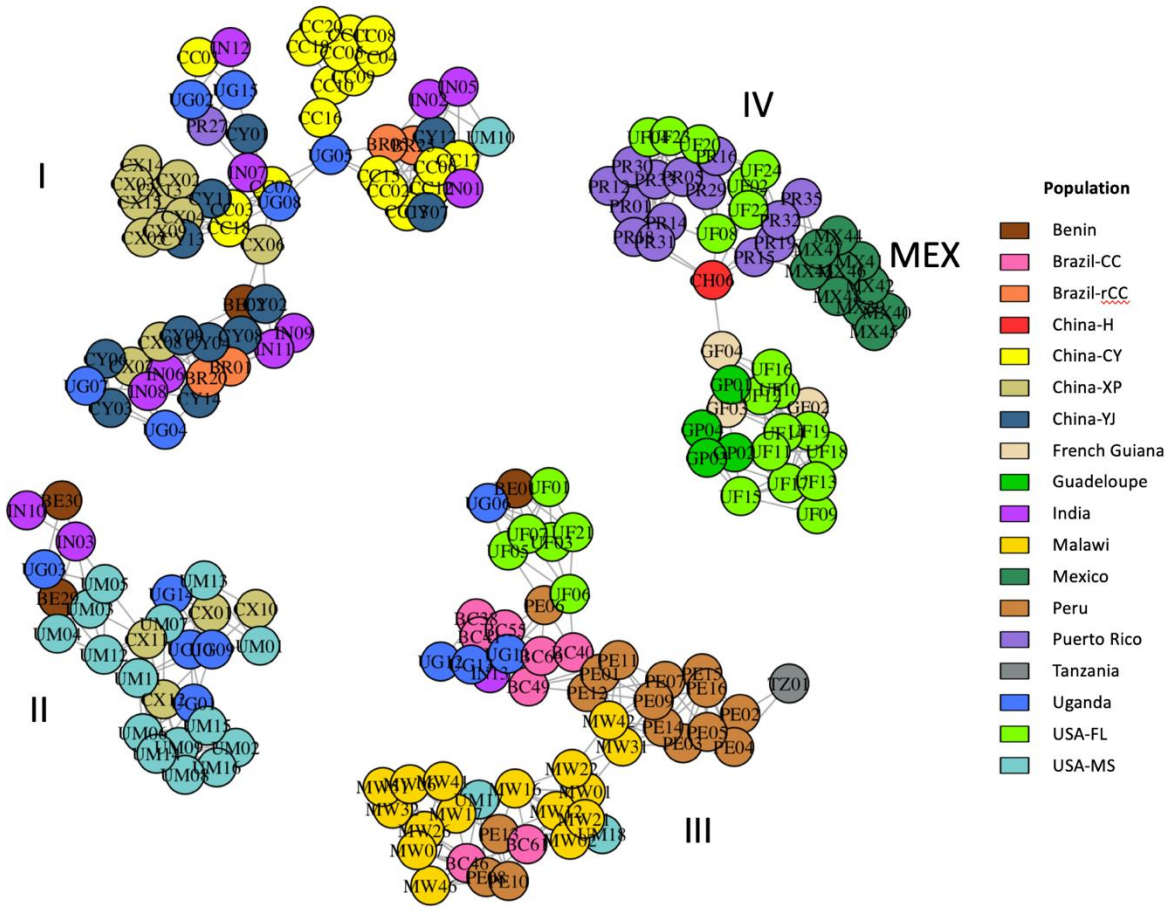
Fig. 5: Admixture analysis based on K=3 (Cross-Validation Error (CVE): 0.375), K=4 (CVE: 0.377), and K=5 (CVE: 0.380). Populations 'USAM' and 'USAF' are from Mississippi and Florida, respectively. Populations from China were from Cangyuan (CY), Yuanjiang (YJ), Xinping (XP) in Yunnan Province. Corn- ('C') or rice- ('R') preferred plant hosts are identified based on mtCOI as per Dumas et al.[55], and by *Tpi* SNP approach as per Nagoshi [62] are indicated by green or yellow bars, respectively. Specimen ID's and sampling countries are as labelled. 'N/C' for *Tpi* indicates no coverage. Grey bars indicate individuals with heterozygous *Tpi* SNPs.



Admixture networks

To explore the population substructure revealed by the admixture analysis in relation to the ML clusters obtained from phylogeny and PCA, we performed network analysis using the plotAdmixture function in the NetView R package. The ML network of individuals belonging to each of the specified populations is shown in Fig. 6A. The four major clusters, I - IV, correspond to those shown in the ML tree (Fig. 3). Individuals from some populations were shown to be spread across multiple clades, e.g., PR, UF and UM from the native range and IN, BE and CX from the invasive populations. Of the populations in the invasive range, those from China were found predominantly in cluster I, with some CX individuals in cluster II and the single CH06 individual in cluster IV.

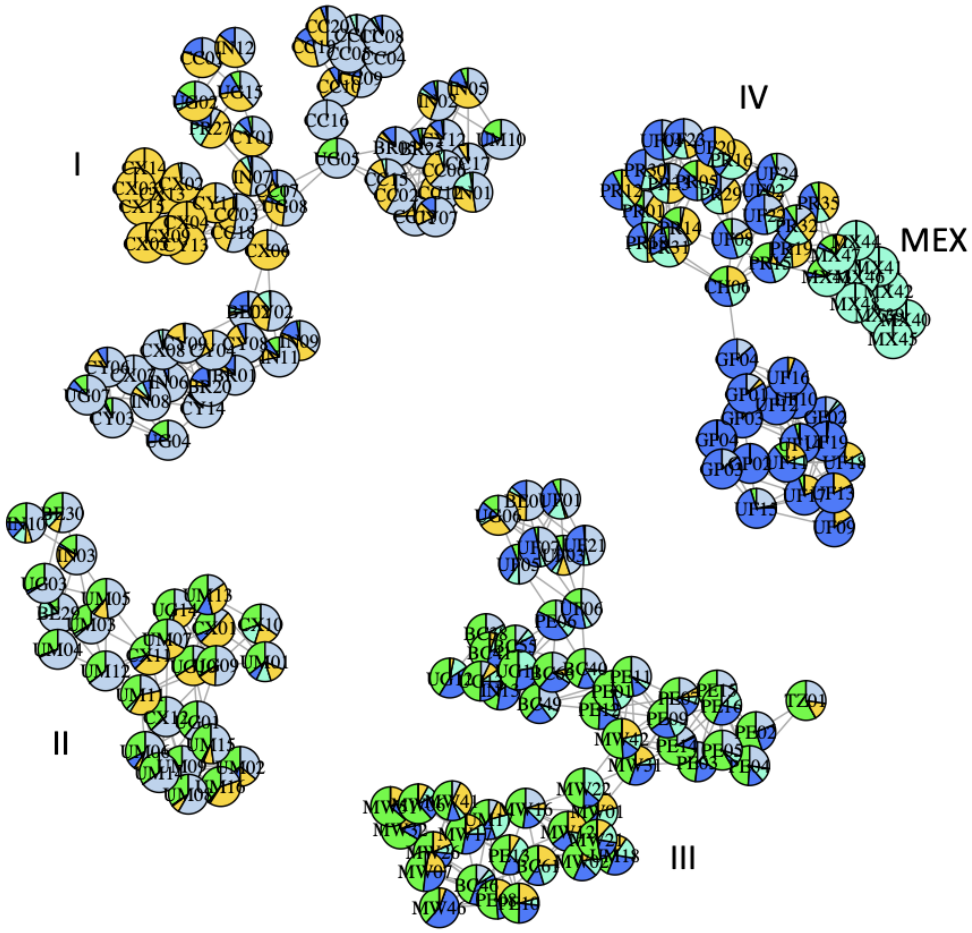
Fig. 6a: Maximum Likelihood (ML) network showing individuals belonging to the populations as given in the legend. The network was drawn using the plotAdmixture function in the R package NetView [97,98], and is based on a ML distance matrix calculated from the IQ-Tree shown in Fig. 3. using the R package ape [99]. The four major clusters, I – IV, correspond to those shown in the tree. Individuals are identified by country codes as follows: China XP (CX), China YJ (CY), China CY (CC), China HIA006 (CH), India (IN), Uganda (UG), Tanzania (TZ), Malawi (MW), Benin (BE), Brazil CC (BC), Brazil rCC (BR), Peru (PE), French Guiana (GF), Mexico (MX), Guadeloupe (GP), Puerto Rico (PR), USA-Florida (UF), and USA-Mississippi (UM). See Supplementary Table 1 for complete information about the individuals.



Plotting admixture proportions at K=5 on this network provided further insight into the complex relationships between these populations (Fig. 6B). The different populations from China that predominantly comprise Cluster I each have distinct admixture profiles that are significantly shared with those of individuals from Uganda and India. In cluster II, China-XP (CX), India (IN), Benin (BE), and Uganda (UG) formed network with

USA-Mississippi (UM) individuals. In Cluster III, all Malawi (MW) individuals and various Ugandan (UG) individuals and the single Tanzanian (TZ) individual formed a network cluster with Peru (PE), Brazil-CC (BC), and some USA Florida (UF) individuals. In cluster IV, only one Chinese FAW (CH) was found to group to this predominantly Caribbean/Central America FAW group (consisting of UF, Puerto Rico (PR), French Guiana (GF), Guadeloupe (GP), and Mexico (MX) FAW individuals).

Fig. 6b: Maximum Likelihood-distance network with admixture analysis at K=5 presented as pie charts for each individual. The ML network is that shown in Figure 6b and individuals, identified by the same codes, and the same four distinct clusters (I-IV) are indicated. Cluster I comprises predominantly different Chinese populations each with distinct admixture profiles but included also individuals from Uganda, India, Brazil-rCC (BR) and Puerto Rico. In cluster II, China-XP (CX), India, Benin, and Uganda formed network with USA-Mississippi individuals. In Cluster III, all Malawi individuals and various Tanzania and Uganda individuals were grouped with Peru, Brazil-CC (BC), and selected USA-FL individuals. In cluster IV, only one Chinese FAW (CH) was found to group to this predominantly Caribbean/Central America FAW group (consisting of USA-FL, Puerto Rico, French Guiana, Guadeloupe, and Mexico FAW individuals). Note that individuals sharing the same colour schemes do not necessarily have the same genetic content, and that the MEX group consisted only of individuals from Mexico showing little admixture with any other population.

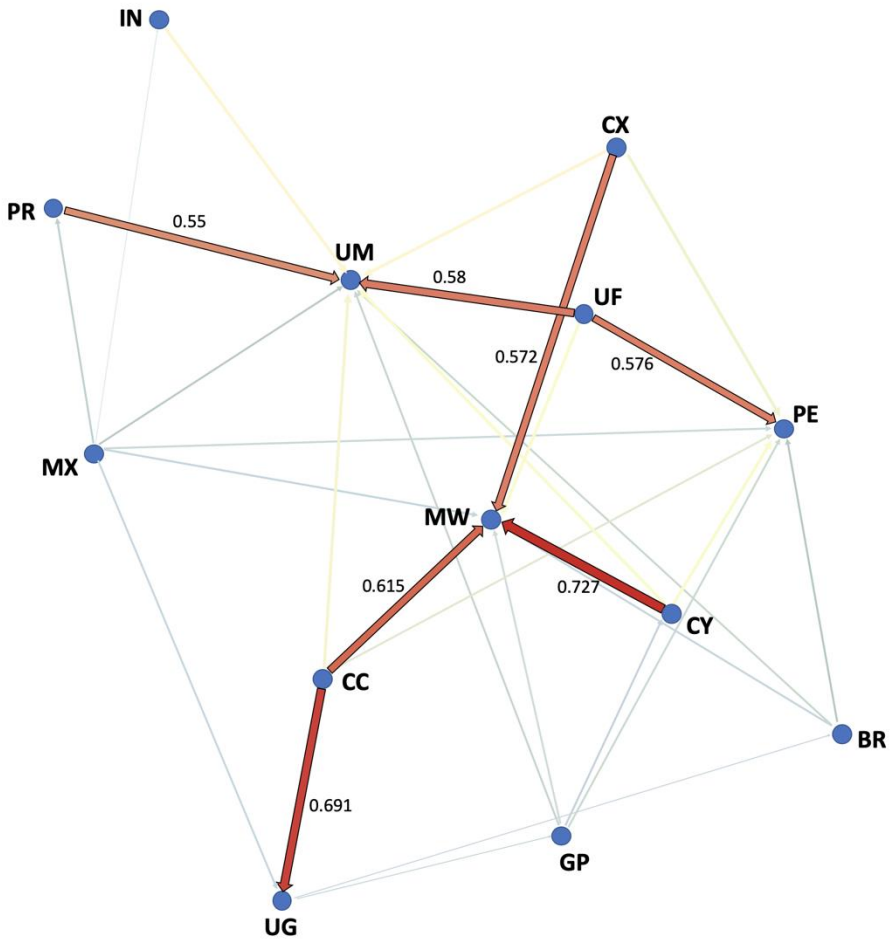


Directionality of gene flow analysis using divMigrate

Analysis of the directionality of gene flow (i.e., relative directional migration) between populations using divMigrate enabled investigation of possible introduction pathways leading to the complex population substructure patterns seen in the above analyses. As shown in Fig. 7, the most significant directional gene flow signatures seen were from all three Chinese populations (i.e., CX, CY, CC) into Malawi and from the Cangyuan

(CC) population into Uganda. Significant gene flow from Florida (UF) and from Puerto Rico (PR) into the Mississippi (UM) FAW population, which the above (e.g., Figs. 3, 4A, 5, 6a) had shown to be distinct was also detected. No evidence was found for directional gene flow from any of the populations studied into China, nor any from or into India. Together with the Admixture results (Fig. 5), these results indicate the East African FAW populations likely originated from China, with some independent 'non-China' introductions also detected in Malawi. The Admixture signatures within the Ugandan FAW population suggested the presence of two genetically distinct FAW populations (Figs. 5, 6B), one of which originated from Asia and involved genetic contribution from the Yunnan Cangyuan (CC) population (Fig. 7), as well as gene flow from Malawi (Fig. 5). While the Malawi population overall showed admixture patterns similar to Peru (Fig. 5) with the PCA showing the Malawi, Peru and Brazil-CC (BC) populations clustered together (Figs. 4B, 4C), directionality analysis indicated genetic contributions from all three Chinese FAW populations (Fig. 7).

Fig 7: Analysis using divMigrate to infer directionality of gene flow (i.e., relative directional migration) between New World native and Old World invasive *Spodoptera frugiperda* populations. The divMigrate analysis was run using the online server <https://popgen.shinyapps.io/divMigrate-online/> [99]. The analysis was performed with the G_{ST} migration statistic of [104] and [105] at filter threshold = 3.0 and 1,000 bootstrap replications to assess confidence with alpha value set at 0.05 (i.e., 95% confidence). Weighted values above 0.50 are indicated. Population codes are IN (India), PR (Puerto Rico), MX (Mexico), UG (Uganda), CC (China Cangyuan), CY (China Yuanjiang), CX (China Xinping), MW (Malawi), PE (Peru), GP (Guadeloupe), BR (Brazil-rCC), UM (USA Mississippi), and UF (USA Florida). Significant gene flow is seen from all three Chinese populations into Malawi and from Cangyuan (CC) to Uganda (UG). Significant gene flow from Florida and from Puerto Rico into the Mississippi FAW population is also detected.



Discussion

The genomic analysis of FAW from native and invasive ranges in this work contradicts recent published theories on the pathway, origin, and direction of spread of this pest across the Old World. Using neutral and unlinked genome-wide SNPs obtained from material available at early stages of the invasion, we showed, through population admixture analysis, ML distance network, and gene flow directionality analyses, that there were likely multiple introductions to both Africa and Asia. Studies to date have relied on analyses of limited partial mitochondrial DNA (e.g., partial COI and CYTB; [21,60]) and the nuclear *Tpi* partial gene (e.g., Nagoshi et al.[59]) of various African, Asian and South East Asian invasive *S. frugiperda* populations, with comparisons to native New World *S. frugiperda* populations. These studies inferred the directionality of spread from the timing of official reporting to the FAO, and described a single introduction of FAW to the Old World from an eastern American/Greater Antilles population, that spread rapidly across the sub-Saharan African nations, before moving to the Indian sub-continent via the Middle East, and then to South East Asia, and China [59].

Genome wide SNP analyses in this present study showed the populations in China and Africa to be genetically diverse and demonstrates strong evidence for a more complex pattern of spread in the Old World, including a significant proportion of east-to-west movement, with populations from Asia (e.g., China Yunnan Province) as a source of invasive FAW populations in Africa (e.g., Malawi, Uganda). The confirmation of FAW after reports by farmers of crop damage, i.e., in Nigeria and São Tomé and Príncipe in early 2016 [17] and in northern and eastern Uganda since 2014 [38], suggested that *S. frugiperda* was present in the African continent earlier, and given the genomic evidence reported here would suggest that the FAW was present in Asia and/or Southeast Asia prior to 2016. Incidences of *S. frugiperda* attacking stem and leaf parts were reported from a farm producing turf grass for parks in Hà Nội, Viet Nam, between March and June (Spring/Summer seasons) of 2008, with heavy infestations reported between the months of April and May in 2008 (see Table 4.2 and Fig. 9 of Vu [106]; see also Nguyen and Vu [107]; Pham [108]), as well as being reported as an agriculture insect pest in the production areas around Hà Nội since 2008 [109]. We also provided clear evidence for multiple introductions of this agricultural pest into Africa, demonstrating conclusively that the Malawian FAW populations have a distinct genomic signature representing a New World population different to the potential New World source populations of the Chinese populations. The early detections and report of FAW in Asia/S.E. Asia (e.g., CH06 from this study from 2016, see GenBank record MT897262; 2008 FAW outbreaks on turf grass in Hà Nội [106]), and the complex pattern of multiple introductions as evident from genomic signatures (e.g., this study; Yainna et al. [69]), are consistent with the perceived rapid spread experienced across the African [110] and Asian continents [111].

Despite being one of the worst agricultural pests in the New World, there has been limited work, at the population genomic level, on the FAW pest complex. Through our genome-wide SNP analyses, we have identified unexpected complexity in the FAW population structure in the New World. While the mitochondrial genome analysis confirmed the two canonical clades that have long been suggested to define two strains with different host preferences, i.e., corn (*Sfc*) and rice (*Sfr*), the neutral nuclear SNP analyses showed a more complex population structure. FAW populations in the New World could be differentiated into at least five distinct groups that broadly followed the species' geographic distributions, and with no obvious pattern related to host race determination by mitochondrial or *Tpi* markers, providing the first genome-wide support for suggestions that these mitochondrial genomes (and the often associated *Tpi* marker) do not define any real population structure across the native range of FAW [57,112], while Groot et al. [113] observed a lack of consistent correlation between host plant and mitochondrial genome in native range populations. Frequent hybridisation has been known to occur in the field (e.g., Nagoshi et al. [114]), and would also account for the observed pattern. Furthermore, Haenniger et al. [115] observed that African populations contained hybrids that were F2 or even later generations, and mating time differences within the African populations were likely related to the differences in circadian gene expression previously identified in *Sfc* or *Sfr* populations in their native range [116]. Orsucci et al. [67] suggested that differences in mitochondria function could be directly related to host preferences, which could explain the absence of a correlation between the mitochondrial and nuclear genotypes, but this lack of genomic correlation points clearly to the need of genome-wide studies in field populations.

In the North American continent, we detected directional migration from Florida and the Puerto Rican populations to the genetically distinct Mississippi one. This is consistent with findings based on mtCOI sequences that the Mississippi populations were established through seasonal migration from Texas and Florida [102]. There also seems to be evidence for a wider Caribbean population including Florida, Puerto Rico, Mexico, the Lesser Antilles (e.g., Guadeloupe) and the north-eastern region of South America (e.g., French Guiana). Mexican FAW formed a separate sub-clade within the Florida/Greater Antilles/Lesser Antilles FAW group. Significant pairwise F_{ST} estimates between Mexico and all New World FAW populations indicated very limited gene flow occurred

between the Mexican population and other New World (and invasive) populations. Northern Mexican populations have been shown to be similar to the Southern Texas overwintering population so it is interesting that in this study the Mexican population sits within the broader Caribbean clade that includes Florida [102]. Previous studies [117,118], for example, have also identified Mexican FAW populations as potentially limited in migratory interactions and biologically unique even between populations from different Mexican geographical regions.

Our PCA on genome wide SNPs identified the Brazilian FAW as two genetically-distinct populations, with one population ('BC') being phylogenetically more closely related to the Peruvian FAW population, and the BR population which is phylogenetically more closely related to the Mississippi population. The Brazilian 'BR' population included individuals that had been found to have a novel 12 bp deletion mutation in the ABCC2 gene [54]. The implications of the close phylogenetic relationship between the BR and Mississippi populations are significant given that FAW is regarded as a major agricultural pest in Brazil [54], and the possible movements of alleles that could potentially underpin resistance, especially to Cry1F and Cry1A toxins, would add to the challenge of managing this pest in the Americas.

The genomic analyses in the present study support multiple introductions from different sources into Africa, rather than via a single introduction that first went into western Africa. Phylogenetic inference and PCA clearly identified the South American FAW population, as represented by the Peru/Brazil (BC) samples, as the likely source for the Malawi population. If this introduction was trade-related, it is more likely to have occurred via neighbouring countries such as South Africa which has greater agricultural trade with South America than does Malawi <https://oec.world/en/> (accessed 06-Dec-2019). A relationship between FAW populations in Uganda, Tanzania and Malawi is evident, with the admixture analysis identifying the Tanzanian individual and selected Ugandan individuals as sharing very similar genomic profiles with the Malawi population (e.g., Fig. 5, K=5,). Interestingly, one Indian FAW individual (i.e., individual IN13; Fig. 5) also shared very similar genomic admixture profiles with the Malawi FAW population, and could indicate movements of East African FAW population into the Asian continent, although it is currently not possible to rule out separate introductions involving the South American Peru/Brazil population into the Indian sub-continent based on limited samples from India.

The complex genetic relationships between populations in Africa and those elsewhere in the Old World have implications for our understanding of pathways for biological invasions. The East African FAW populations from Uganda and Malawi showed genetic relatedness to the Yunnan CC population, and the Yunnan CC, CY, and CX populations, respectively, with directionality analyses showing gene flow from China to Africa, consistent with admixture analyses, although we also note that our inference of FAW movements across the Old World is significantly limited by the availability of Asian/East Asian/South East Asian populations. The ML network analysis showed that the Benin FAW population also clustered within the same networks as the Yunnan CY, CX and Ugandan populations (Clusters I and II), indicating that the spread of FAW from Yunnan into the African continent potentially extended to west Africa. The detection of a shared rice mitogenome haplotype between the two extremes of our invasive FAW sampling sites of Benin, and from the Yunnan CY population from China, provided further evidence for this 'East to West' spread of FAW populations into Africa (Fig. 1). While it might be logical to conclude, from the very limited data of shared mitogenome haplotypes between Benin and China, that FAW spread was from west Africa to east Asia, applying bioinformatic analytical approaches enabled a re-evaluation of the FAW global invasion patterns. For example, copy number variations (CNV) analysis from whole genome sequence data [38] also detected distinct signatures between the China/East African (Uganda, Malawi) from the Benin/Indian FAW populations, thereby further supported multiple introductions of FAW across different Old World regions. This and the study of Yainna et al. [69] highlight the need for whole-genome derived datasets to accurately trace population origins and global movements of highly mobile insect pests.

The invasive FAW in west Africa (i.e., Togo) was proposed to have originated from the Greater Antilles/Eastern Florida based on inferences from a partial nuclear *Tpi* gene marker and two partial mitochondrial DNA genes (i.e., COI, Cyt *b*) [19], and which was further linked to subsequent invasion into South-eastern Asia and China [59]. This finding was not supported by our whole genome analysis, which identified at least four distinct populations of FAW in Yunnan Province, involving at least five separate introduction events (Figs. 6a, 6b). While the Yunnan CH06 individual and the Yunnan CX, CY and CC populations all possessed their own unique genetic signatures, only the various CX individuals could be linked to the USA Mississippi FAW population, and the single CH06 FAW individual (Figs. 3 and 4F) to the USA Florida population. The Yunnan CY and CC populations and the remaining CX individuals all clustered with the Brazil-rCC (i.e., 'BR') population (Figs. 3 and

4F), potentially indicating a shared origin, or the same source population and as a connected introduction event across globally disparate locations.

While populations in the Yunnan Province and Africa have enabled inference of the overall directionality of introductions, the lack of samples across much of the invasive range, in particular S.E. Asia (e.g., Myanmar, Thailand, Vietnam, Malaysia, Indonesia, etc.) meant that other candidate FAW invasive populations could have been the 'source invasive populations' for the Asian/Old World invasions. One example is the modelling of FAW spread via monsoon wind patterns from Myanmar into southern China [119], a hypothesis that could be tested using genomic evidence. International trade pathways are increasingly being identified as responsible for accidental introductions of invasive plant pests and pathogens (e.g., Lopes-da-Silva [120]). Evaluation of introduction pathways will therefore need to also include trade data, as has been undertaken for the invasion by *H. armigera* from the Old World into the New World [3,15] (reviewed also in Jones et al. [9]), and for the invasive *Bemisia tabaci* MED and MEAM1 species complex (reviewed by De Barro et al. [2]; see also Elfekih et al. [1]).

Indeed, global movements of invasive pests, exemplified by the spread of FAW, as well as the multiple introductions of *H. armigera* into the South American continent [15] and various Caribbean nations [121] from Asia [10,14] and various Old World regions [3], are timely reminders of the need for global coordination of enhanced biosecurity preparedness strategies that build on advancement in genomic research. The potential negative impacts of introductions of alien species include introgression of genetic traits to local species through hybridisation [10,13,14,16,28]. Development of new trans-continental trade routes to increase economic growth between trading partners must therefore recognise the significant risks and take into consideration the biosecurity implications associated with the rapid spreading of highly invasive pests and pathogens of plants, animals and humans [122] that could instead undermine the aim to grow the global economy.

Acknowledgements

This work was funded by CSIRO Health & Biosecurity (R-91044-01) to WTT, TW, KHJG, SE, SD. WTT, TW, SE, KHJ, DK were also funded by CSIRO Health & Biosecurity 'Genes of Biosecurity importance fund' (R-8681-1), and DK also acknowledged funding by CSIRO (R-90035-14). RR was funded by CSIRO and Hort Innovation Australia (HIA). AP was supported by CSIRO, CRDC and Bayer (R10801-01). JZ and YW were supported by National Key Research Development Program of China (No. 2019YFD0300103 to YW) and the Fundamental Research Funds for the Central Universities of China (KYZ201920 to YW). MHO was funded by the Agricultural Technology and Agribusiness Advisory Services Project (Project ID No.: P109224) funded by the World Bank. The work performed at INRAE was publicly funded through the ANR (the French National Research Agency, Grant ID 1702-018, given to KN) under the "Investissements d'avenir" programme with the reference ANR-10-LABX-001-01 Labex Agro and coordinated by Agropolis Fondation under the frame of I-SITE MUSE (ANR-16-IDEX-0006). It was also funded by a grant from the department of Santé des Plantes et Environnement at Institut national de la recherche agronomique for KN (adaptivesv). It was also financially supported by EUPHRESCO (FAW-spedcom, given to Anne-Nathalie Volkoff). We thank Stella Adumo (NaCRRI) with Uganda FAW sample collection. Peru and China FAW pre-border interception specimens were provided by the DAWE. We thank Paul De Barro and Andy Sheppard (CSIRO) for helpful discussion. Dr Dao Thi Hang (Plant Protection Research Institute, Vietnam) assisted with sourcing and translation of reference material from Vietnam National University of Agriculture. Dr Nguyen Van Liem, Dr Dao Thi Hang (Plant Protection Research Institute, Vietnam), and Prof. Dr. Vu Van Lien (Deputy General Director, Vietnam National Museum of Nature, VAST) for helpful discussion.

Statement on Data Availability

All assembled mitochondrial genomes have been submitted to GenBank (accession numbers MT897262 - MT897458). The complete list of FAW population genome wide single nuclear polymorphic loci used is available from CSIRO's public data access portal (doi: xxxxxxxxxx).

Statement on Code Availability

All programme codes used in this study have been cited. No custom program codes were used in this study.

References

1 Elfekih, S. *et al.* Genome-wide analyses of the *Bemisia tabaci* species complex reveal contrasting patterns of admixture and complex demographic histories. *PLoS One* **13**, e0190555, doi:10.1371/journal.pone.0190555 (2018).

2 De Barro, P. J., Liu, S. S., Boykin, L. M. & Dinsdale, A. B. *Bemisia tabaci*: a statement of species status. *Annu Rev Entomol* **56**, 1-19, doi:10.1146/annurev-ento-112408-085504 (2011).

3 Tay, W. T. *et al.* Mitochondrial DNA and trade data support multiple origins of *Helicoverpa armigera* (Lepidoptera, Noctuidae) in Brazil. *Sci Rep* **7**, 45302, doi:10.1038/srep45302 (2017).

4 Czepak, C., Albernaz, C., Vivan, L. M., Guimarães, H. O. & Carvalhais, T. First reported occurrence of *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) in Brazil. *Pesq. Agropec. Trop., Goiânia* **43**, 110-113 (2013).

5 Tay, W. T. *et al.* A Brave New World for an Old World Pest: *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Brazil. *Plos One* **8**, doi: 10.1371/journal.pone.0080134 (2013).

6 Gilligan, T. M. *et al.* Identification of Heliothine (Lepidoptera: Noctuidae) Larvae Intercepted at U.S. Ports of Entry From the New World. *J Econ Entomol* **112**, 603-615, doi:10.1093/jee/toy402 (2019).

7 Kriticos, D. J. *et al.* Correction: The Potential Distribution of Invading *Helicoverpa armigera* in North America: Is It Just a Matter of Time? *PLoS One* **10**, e0133224, doi:10.1371/journal.pone.0133224 (2015).

8 Kriticos, D. J. *et al.* The potential distribution of invading *Helicoverpa armigera* in North America: is it just a matter of time? *PLoS One* **10**, e0119618, doi:10.1371/journal.pone.0119618 (2015).

9 Jones, C. M., Parry, H., Tay, W. T., Reynolds, D. R. & Chapman, J. W. Movement Ecology of Pest *Helicoverpa*: Implications for Ongoing Spread. *Annu Rev Entomol* **64**, 277-295, doi:10.1146/annurev-ento-011118-111959 (2019).

10 Walsh, T. K. *et al.* Multiple recombination events between two cytochrome P450 loci contribute to global pyrethroid resistance in *Helicoverpa armigera*. *PLoS One* **13**, e0197760, doi:10.1371/journal.pone.0197760 (2018).

11 Pomari-Fernandes, A., de Freitas Bueno, A. & Sosa-Gomez, D. R. *Helicoverpa armigera*: current status and future perspectives in Brazil. . *Current Agricultural Science and Technology* **21**, 1-7 (2015).

12 Pozebon, H. *et al.* Arthropod Invasions Versus Soybean Production in Brazil: A Review. *J Econ Entomol*, doi:10.1093/jee/toaa108 (2020).

13 Anderson, C. J. *et al.* Hybridization and gene flow in the mega-pest lineage of moth, *Helicoverpa*. *Proc Natl Acad Sci U S A* **115**, 5034-5039, doi:10.1073/pnas.1718831115 (2018).

14 Anderson, C. J., Tay, W. T., McGaughran, A., Gordon, K. & Walsh, T. K. Population structure and gene flow in the global pest, *Helicoverpa armigera*. *Mol Ecol* **25**, 5296-5311, doi:10.1111/mec.13841 (2016).

15 Arnemann, J. A. *et al.* Multiple incursion pathways for *Helicoverpa armigera* in Brazil show its genetic diversity spreading in a connected world. *Sci Rep* **9**, 19380, doi:10.1038/s41598-019-55919-9 (2019).

16 Valencia-Montoya, W. A. *et al.* Adaptive introgression across semipermeable species boundaries between local *Helicoverpa zea* and invasive *Helicoverpa armigera* moths. *Mol Biol Evol*, doi:10.1093/molbev/msaa108 (2020).

17 Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamo, M. 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**, e0165632, doi:10.1371/journal.pone.0165632 (2016).

18 Cock, M. J. W., Beseh, P. K., Buddie, A. G., Cafa, G. & Crozier, J. Molecular methods to detect *Spodoptera frugiperda* in Ghana, and implications for monitoring the spread of invasive species in developing countries. *Sci Rep* **7**, 4103, doi:10.1038/s41598-017-04238-y (2017).

19 Nagoshi, R. N. *et al.* 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**, e0181982, doi:10.1371/journal.pone.0181982 (2017).

20 Jacobs, A., van Vuuren, A. & Rong, I. H. Characterisation of the fall armyworm (*Spodoptera frugiperda* JE Smith) (Lepidoptera: Noctuidae) from South Africa. *Afr Entomol* **26**, 45-49, doi:10.4001/003.026.0045 (2018).

21 Otim, M. H. *et al.* Detection of sister-species in invasive populations of the fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae) from Uganda. *PLoS One* **13**, e0194571, doi:10.1371/journal.pone.0194571 (2018).

22 FAO. *Briefing note on FAO actions on fall armyworm in Africa*, http://www.fao.org/3/a-bt415e.pdf (2018).

23 FAO. *Briefing note on FAO actions on fall armyworm*, http://www.fao.org/3/BS183E/bs183e.pdf (2019).

24 Ganiger, P. C. *et al.* Occurrence of the new invasive pest, fall armyworm, *Spodoptera frugiperda* (JE Smith) (Lepidoptera: Noctuidae), in the maize fields of Karnataka, India. *Curr Sci India* **115**, 621-623, doi: 10.18520/cs/v115/i4/621-623 (2018).
25 Sharanabasappa, D. *et al.* First report of the fall Armyworm, *Spodoptera frugiperda* (J E Smith) (Lepidoptera, Noctuidae) an Alien invasive pest on Maize in India. *Pest management in Horticultural Ecosystems* **24**, 23-29 (2018).
26 FAO. in *FAO Regional Conference for Asia and the Pacific, 35th Session* APRC/20/7 (Thimphu, Bhutan, 2019).
27 EPPO. First report of Spodoptera frugiperda in Thailand. (2019).
28 Tay, W. T. & Gordon, K. H. J. Going global - genomic insights into insect invasions. *Curr Opin Insect Sci* **31**, 123-130, doi:10.1016/j.cois.2018.12.002 (2019).
29 Zhang, L. *et al.* Molecular identification of invasive fall armyworm *Spodoptera frugiperda* in Yunnan Province. *Plant Protection* **45**, 19-24 (2019).
30 Wu Q, Jian Y & K., W. Analysis of migration routes of the fall armyworm *Spodoptera frugiperda* (J. E. Smith) form Myanmar to China. *Plant Protection* **45**, 1-6 (2019).
31 USDA. Fall armyworm damages corn and threatens other crops in Vietnam. 7 (United States Department of Agriculture, Foreign Agricultural Serves, 2019).
32 FAO. Report of first detection of fall armyworm (FAW) in the Republic of the Philippines. Report No. PHL-02/1, (Food and Agriculture Organization of the United Nations, International Plant Protection Convention, 2019).
33 Navasero, M. V. *et al.* Detection of the fall armyworm, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera: Noctuidae) using larval mrophological characters, and observations on its current local distribution in the Philippines. *Philipp Ent* **33**, 171-184 (2019).
34 Vennila, S. *et al.* in *International Workshop on Facilitating International Research Collaboration on Transboundary Plant Pests.* 34 (Ministry of Agriculture, Forestry and Fisheries, Tsukuba, Ibaraki, Japan, 2019).
35 FAO. First detection of fall armyworm in China. (Food and Agriculture Organization of the United Nations, International Plant Protection Convention, 2019).
36 Silver, A. Caterpillar's devastating march across China spurs hunt for native predator. *Nature* **570**, 286-287, doi:10.1038/d41586-019-01867-3 (2019).
37 Song, X. P. *et al.* Intrusion of Fall Armyworm (*Spodoptera frugiperda*) in Sugarcane and Its Control by Drone in China. *Sugar Tech* **22**, 734-737, doi:10.1007/s12355-020-00799-x (2020).
38 Czepak, C. *et al.* Especial Spodoptera: Migração acelerada. *Cultivar Grandes Culturas* **244**, 26-29 (2019).
39 FAO. First detection of Fall armyworm in Torres Strait of Australia. (Food and Agriculture Organization of the United Nations, International Plant Protection Convention, 2020).
40 Queensland Government, Department of Agriculture and Fisheries. First mainland detection of fall armyworm. News Release: 19-Feb-2020. (2020).
41 Wild, S. Invasive pest hits Africa. *Nature* **543**, 13-14, doi: 10.1038/nature.2017.21527 (2017).
42 Early, R., Gonzalez-Moreno, P., Murphy, S. T. & Day, R. Forecasting the global extent of invasion of the cereal pest *Spodoptera frugiperda*, the fall armyworm. *Neobiota*, 25-50, doi:10.3897/neobiota.40.28165 (2018).
43 FAO. Fall armyworm likely to spread from India to other parts of Asia with South East Asia and South China most at risk. (Food and Agriculture Organization of the United Nation, 2018).
44 Gouin, A. *et al.* Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. *Sci Rep* **7**, 11816, doi:10.1038/s41598-017-10461-4 (2017).
45 Zhang, L. *et al.* Genetic structure and insecticide resistance characteristics of fall armyworm populations invading China. *Mol Ecol Resour*, doi:10.1111/1755-0998.13219 (2020).
46 Westbrook, J., Fleischer, S., Jairam, S., Meagher, R. & Nagoshi, R. Multigenerational migration of fall armyworm, a pest insect. *Ecosphere* **10**, e02919 doi: 10.1002/ecs2.2919 (2019).
47 du Plessis, H., van den Berg, J., Ota, N. & Kriticos, D. J. in *CSIRO-InSTePP Pest Geography.* 7 (CSIRO).
48 FAO. First detection report of the fall armyworm *Spodoptera frugiperda* (Lepdioptera: Noctuidae) on maize in Myanmar., (Food and Agriculture Organization of the United Nations, International Plant Protection Convention, 2019).
49 Sun, X.-x. *et al.* Case study on the first immigration of fall armyworm *Spodoptera frugiperda* invading into China. *Journal of Integrative Agriculture* **18**, 2-10, DOI: 10.1016/S2095-3119(19)62839-X. (2019).
50 Day, R. *et al.* Fall armyworm: impacts and implications for Africa. *Outlooks on Pest Management* **28**, 196-201 DOI: https://doi.org/10.1564/v28_oct_02 (2017).

51 Assefa, F. & Ayalew, D. Status and control measures of fall armyworm (*Spodoptera frugiperda*) infestations in maize fields in Ethiopia: A review. *Cogent Food Agr* **5**, doi:Unsp 1641902 10.1080/23311932.2019.1641902 (2019).
52 Hurska, A. J. Fall armyworm (*Spodoptera frugiperda*) management by smallholders. *CAB Reviews* **14**, 11, doi: 10.1079/PAVSNNR201914043 (2019).
53 Firake, D. M. & Behere, G. T. Natural mortality of invasive fall armyworm, *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae) in maize agroecosystems of northeast India. *Biological Control* **148**, 104303, doi:https://doi.org/10.1016/j.biocontrol.2020.104303 (2020).
54 Guan, F. *et al.* Whole-genome sequencing to detect mutations associated with resistance to insecticides and Bt proteins in *Spodoptera frugiperda*. *Insect Sci*, doi:10.1111/1744-7917.12838 (2020).
55 Dumas, P. *et al.* Phylogenetic molecular species delimitations unravel potential new species in the pest genus *Spodoptera* Guenee, 1852 (Lepidoptera, Noctuidae). *PLoS One* **10**, e0122407, doi:10.1371/journal.pone.0122407 (2015).
56 Dumas, P. *et al. Spodoptera frugiperda* (Lepidoptera: Noctuidae) host-plant variants: two host strains or two distinct species? *Genetica* **143**, 305-316, doi:10.1007/s10709-015-9829-2 (2015).
57 Nagoshi, R. N. *et al.* Genetic characterization of fall armyworm (*Spodoptera frugiperda*) in Ecuador and comparisons with regional populations identify likely migratory relationships. *PLoS One* **14**, e0222332, doi:10.1371/journal.pone.0222332 (2019).
58 Jing, D. P. *et al.* Initial detections and spread of invasive *Spodoptera frugiperda* in China and comparisons with other noctuid larvae in cornfields using molecular techniques. *Insect Sci* **27**, 780-790, doi:10.1111/1744-7917.12700 (2020).
59 Nagoshi, R. N. *et al.* Southeastern Asia fall armyworms are closely related to populations in Africa and India, consistent with common origin and recent migration. *Sci Rep* **10**, 1421, doi:10.1038/s41598-020-58249-3 (2020).
60 Mahadeva, S. H. M. *et al.* Prevalence of "R" strain and molecular diversity of fall army worm *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera: Noctuidae) in India. *Indian J Entomol* **80**, 544-553, doi:10.5958/0974-8172.2018.00239.0 (2018).
61 Murua, M. G. *et al.* Demonstration Using Field Collections that Argentina Fall Armyworm Populations Exhibit Strain-specific Host Plant Preferences. *Journal of Economic Entomology* **108**, 2305-2315, doi:10.1093/jee/tov203 (2015).
62 Nagoshi, R. N. The Fall Armyworm Triose Phosphate Isomerase (Tpi) Gene as a Marker of Strain Identity and Interstrain Mating. *Ann Entomol Soc Am* **103**, 283-292, doi:10.1603/An09046 (2010).
63 Nagoshi, R. N., Goergen, G., Plessis, H. D., van den Berg, J. & Meagher, R., Jr. Genetic comparisons of fall armyworm populations from 11 countries spanning sub-Saharan Africa provide insights into strain composition and migratory behaviors. *Sci Rep* **9**, 8311, doi:10.1038/s41598-019-44744-9 (2019).
64 Behere, G. T. *et al.* Mitochondrial DNA analysis of field populations of *Helicoverpa armigera* (Lepidoptera: Noctuidae) and of its relationship to H. zea. *BMC Evol Biol* **7**, 117, doi:10.1186/1471-2148-7-117 (2007).
65 Pearce, S. L. *et al.* Erratum to: Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC Biol* **15**, 69, doi:10.1186/s12915-017-0413-3 (2017).
66 Pearce, S. L. *et al.* Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC Biol* **15**, 63, doi:10.1186/s12915-017-0402-6 (2017).
67 Orsucci, M. *et al.* Transcriptional plasticity evolution in two strains of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) feeding on alternative host-plants. *bioRxiv*, 263186, doi:10.1101/263186 (2018).
68 Nam, K. *et al.* Adaptation by copy number variation increases insecticide resistance in fall armyworms. *bioRxiv*, 812958, doi:10.1101/812958 (2019).
69 Yainna, S. *et al.* Genomic balancing selection is key to the invasive success of the fall armyworm. *bioRxiv*, 2020.2006.2017.154880, doi:10.1101/2020.06.17.154880 (2020).
70 Tay, W. T. *et al.* Novel molecular approach to define pest species status and tritrophic interactions from historical *Bemisia* specimens. *Sci Rep-Uk* **7**, doi:ARTN 429 10.1038/s41598-017-00528-7 (2017).
71 Walsh, T. K. *et al.* Mitochondrial DNA genomes of five major *Helicoverpa* pest species from the Old and New Worlds (Lepidoptera: Noctuidae). *Ecol Evol* **9**, 2933-2944, doi:10.1002/ece3.4971 (2019).
72 Bernt, M. *et al.* MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* **69**, 313-319, doi:10.1016/j.ympev.2012.08.023 (2013).

73 Villesen, P. FaBox: an online toolbox for FASTA sequences. *Mol Ecol Notes* **7**, 965-968, doi:10.1111/j.1471-8286.2007.01821.x (2007).
74 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066, doi:10.1093/nar/gkf436 (2002).
75 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).
76 Nam, K. *et al.* Divergent selection causes whole genome differentiation without physical linkage among the targets in *Spodoptera frugiperda* (Noctuidae). *bioRxiv*, 452870, doi:10.1101/452870 (2018).
77 Liu, H. *et al.* Chromosome level draft genomes of the fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae), an alien invasive pest in China. *bioRxiv*, 671560, doi:10.1101/671560 (2019).
78 Xiao, H. *et al.* The genetic adaptations of fall armyworm *Spodoptera frugiperda* facilitated its rapid global dispersal and invasion. *Mol Ecol Resour* **20**, 1050-1068, doi:10.1111/1755-0998.13182 (2020).
79 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997v2 [q-bio.GN]* (2013).
80 Bushnell, B. BBMap: A Fast, Accurate, Splice-Aware Aligner., (Lawrence Berkeley National Laboratory. , 2014).
81 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
82 Trifinopoulos, J., Nguyen, L. T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **44**, W232-235, doi:10.1093/nar/gkw256 (2016).
83 Minh, B. Q., Nguyen, M. A. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**, 1188-1195, doi:10.1093/molbev/mst024 (2013).
84 Huson, D. H. & Scornavacca, C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol* **61**, 1061-1067, doi:10.1093/sysbio/sys062 (2012).
85 Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: an analysis tool set for population genomics. *Mol Ecol* **22**, 3124-3140, doi:10.1111/mec.12354 (2013).
86 Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405, doi:10.1093/bioinformatics/btn129 (2008).
87 Jombart, T. & Ahmed, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070-3071, doi:10.1093/bioinformatics/btr521 (2011).
88 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7, doi:10.1186/s13742-015-0047-8 (2015).
89 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158, doi:10.1093/bioinformatics/btr330 (2011).
90 Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595 (1989).
91 Fu, Y. X. & Li, W. H. Statistical tests of neutrality of mutations. *Genetics* **133**, 693-709 (1993).
92 Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol* **31**, 1929-1936, doi:10.1093/molbev/msu136 (2014).
93 Wright, S. The Genetical Structure of Populations. *Ann Eugenic* **15**, 323-354 (1951).
94 Raymond, M. & Rousset, F. Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. *J Hered* **86**, 248-249, doi: 10.1093/oxfordjournals.jhered.a111573 (1995).
95 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).
96 Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664, doi:10.1101/gr.094052.109 (2009).
97 Neuditschko, M., Khatkar, M. S. & Raadsma, H. W. NetView: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. *PLoS One* **7**, e48375, doi:10.1371/journal.pone.0048375 (2012).
98 Steinig, E. J., Neuditschko, M., Khatkar, M. S., Raadsma, H. W. & Zenger, K. R. netview p: a network visualization tool to unravel complex population structure using genome-wide SNPs. *Mol Ecol Resour* **16**, 216-227, doi:10.1111/1755-0998.12442 (2016).
99 Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528, doi:10.1093/bioinformatics/bty633 (2019).
100 Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. & Prodohl, P. A. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* **4**, 782-788, doi:10.1111/2041-210x.12067 (2013).

101 Sundqvist, L., Keenan, K., Zackrisson, M., Prodohl, P. & Kleinhans, D. Directional genetic differentiation and relative migration. *Ecol Evol* **6**, 3461-3475, doi:10.1002/ece3.2096 (2016).
102 Nagoshi, R. N., Meagher, R. L. & Hay-Roe, M. Inferring the annual migration patterns of fall armyworm (Lepidoptera: Noctuidae) in the United States from mitochondrial haplotypes. *Ecol Evol* **2**, 1458-1467, doi:10.1002/ece3.268 (2012).
103 Wright, S. The Interpretation of Population-Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution* **19**, 395-420, doi: 10.2307/2406450 (1965).
104 Nei, M. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A* **70**, 3321-3323, doi:10.1073/pnas.70.12.3321 (1973).
105 Nei, M. & Chesser, R. K. Estimation of fixation indices and gene diversities. *Ann Hum Genet* **47**, 253-259, doi:10.1111/j.1469-1809.1983.tb00993.x (1983).
106 Vu, T. P. *Insect pests of turf grass, biology, ecology and the control of Herpetogramma phaeoptralis (Guenée) in Hà Nội in Spring Summer 2008* PhD thesis, Hà Nội Agriculture University, Vietnam, (2008).
107 Nguyen, T. K. O. & Vu, T. P. in *Scientific Report on Ecology and Biological Resources. The third National Science Conference.* 1490-1498.
108 Pham, V. L. in *Plant Protection Magazine* Vol. No. 4/2019 (Plant Protection Research Institute of Vietnam, 2019).
109 Nguyen, V. D., Ha, Q. H. & Nguyen, T. T. C. in *Vietnam insects and pests* (ed V. L. Pham) (2012).
110 Stokstad, E. FOOD SECURITY New crop pest takes Africa at lightning speed. *Science* **356**, 473-474, doi: 10.1126/science.356.6337.473 (2017).
111 Baloch, M. N., Fan, J. Y., Haseeb, M. & Zhang, R. Z. Mapping Potential Distribution of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Central Asia. *Insects* **11**, 172 doi: 10.3390/insects11030172 (2020).
112 Juarez, M. L. *et al.* Population structure of *Spodoptera frugiperda* maize and rice host forms in South America: are they host strains? *Entomol Exp Appl* **152**, 182-199, doi:10.1111/eea.12215 (2014).
113 Groot, A. T. *et al.* Evolution of Reproductive Isolation of *Spodoptera frugiperda*. *Pheromone Communication in Moths: Evolution, Behavior, and Application*, 291-300 (2016).
114 Nagoshi, R. N., Meagher, R. L., Nuessly, G. & Hall, D. G. Effects of fall armyworm (Lepidoptera : Noctuidae) interstrain mating in wild populations. *Environ Entomol* **35**, 561-568, doi: 10.1603/0046-225x-35.2.561 (2006).
115 Haenniger, S. *et al.* Sexual communication of *Spodoptera frugiperda* from West Africa: Adaptation of an invasive species and implications for pest management. *Sci Rep* **10**, 2892, doi:10.1038/s41598-020-59708-7 (2020).
116 Hanniger, S. *et al.* Genetic basis of allochronic differentiation in the fall armyworm. *Bmc Evolutionary Biology* **17**, 68 doi: 10.1186/s12862-017-0911-5 (2017).
117 Lopez-Edwards, M. *et al.* Biological differences between five populations of fall armyworm (Lepidoptera : Noctuidae) collected from corn in Mexico. *Fla Entomol* **82**, 254-262, doi: 10.2307/3496577 (1999).
118 Nagoshi, R. N. *et al.* Haplotype Profile Comparisons Between *Spodoptera frugiperda* (Lepidoptera: Noctuidae) Populations From Mexico With Those From Puerto Rico, South America, and the United States and Their Implications to Migratory Behavior. *Journal of Economic Entomology* **108**, 135-144, doi:10.1093/jee/tou044 (2015).
119 Wu, Q. L. *et al.* Estimation of the Potential Infestation Area of Newly-invaded Fall Armyworm *Spodoptera frugiperda* in the Yangtze River Valley of China. *Insects* **10**, 298 doi:10.3390/insects10090298 (2019).
120 Lopes-da-Silva, M., Sanches, M. M., Stancioli, A. R., Alves, G. & Sugayama, R. The Role of Natural and Human-Mediated Pathways for Invasive Agricultural Pests: A Historical Analysis of Cases from Brazil. . *Agricultural Sciences* **5**, 634-646, doi:http://dx.doi.org/10.4236/as.2014.57067 (2014).
121 Tembrock, L. R., Timm, A. E., Zink, F. A. & Gilligan, T. M. Phylogeography of the Recent Expansion of *Helicoverpa armigera* (Lepidoptera: Noctuidae) in South America and the Caribbean Basin. *Ann Entomol Soc Am* **112**, 388-401, doi:10.1093/aesa/saz019 (2019).
122 Liu, X. *et al.* Risks of Biological Invasion on the Belt and Road. *Curr Biol* **29**, 499-505.e4, doi:10.1016/j.cub.2018.12.036 (2019).