



# OPEN Invasive fall armyworms are corn strain

Karine Durand<sup>1,2</sup>, Hyerin An<sup>1,2</sup> & Kiwoong Nam<sup>1</sup>✉

The fall armyworm (*Spodoptera frugiperda*) is one of the major pest insects in diverse crop plants, including maize, rice, and cotton. While the fall armyworm is native to North and South America, its invasion was first reported in West Africa in 2016. Since then, this species has rapidly spread across Sub-Saharan Africa, Asia, and Oceania, as well as Egypt and Cyprus. The fall armyworm is composed of two sympatric strains, the corn and rice strains, designated to their preferred host plants, in native areas. It remains surprisingly unclear whether invasive fall armyworms belong to the corn strain, rice strain, or hybrids of the two, despite a large number of population genetics studies. In this study, we performed population genomics analyses using globally collected 116 samples to identify the strains of invasive fall armyworms. We observed that invasive fall armyworms are genomically most similar to the corn strain. The reconstructed phylogenetic tree supports the hypothesis that invasive fall armyworms originated from the corn strain. All genomic loci of invasive populations exhibit higher genetic similarity to the corn strains compared to the rice strains. Furthermore, we found no evidence of gene flow from rice strains to invasive populations at any genomic locus. These results demonstrate that invasive fall armyworms belong to the corn strain. These results suggest that invasive fall armyworms likely have very limited potential to infest rice. Therefore, the management plan should primarily focus on crops preferred by the corn strain.

**Keywords** Corn strain, Fall armyworm, Invasion, Pest genomics, *Spodoptera frugiperda*

The fall armyworm (*Spodoptera frugiperda*; Noctuidae; Lepidoptera, FAW) is one of the major pest insects of diverse crops including cotton, maize, rice, and sorghum. FAW is native to North and South America, and its invasion was first reported in West Africa in 2016<sup>1</sup>. Since then, FAW has rapidly spread to various regions, including Sub-Saharan Africa, South Asia, Southeast Asia, East Asia, Oceania, Egypt, and more recently, Cyprus and the Canary Islands (<https://www.fao.org/fall-armyworm/monitoring-tools/faw-map/en/>). Invasive FAWs cause significant yield losses in maize production. For example, FAW reduced up to 23–53% of maize production in Sub-Saharan Africa<sup>2</sup>, where maize provides at least 30% of total caloric intake<sup>3</sup>. Field-evolved resistance to insecticides appears to be widespread<sup>4–9</sup>, posing challenges for FAW control. Given the global and urgent nature of the FAW invasion, it is crucial to monitor and assess the negative impacts of invasive FAWs, as well as to develop sustainable and ecologically sound methods for FAW control.

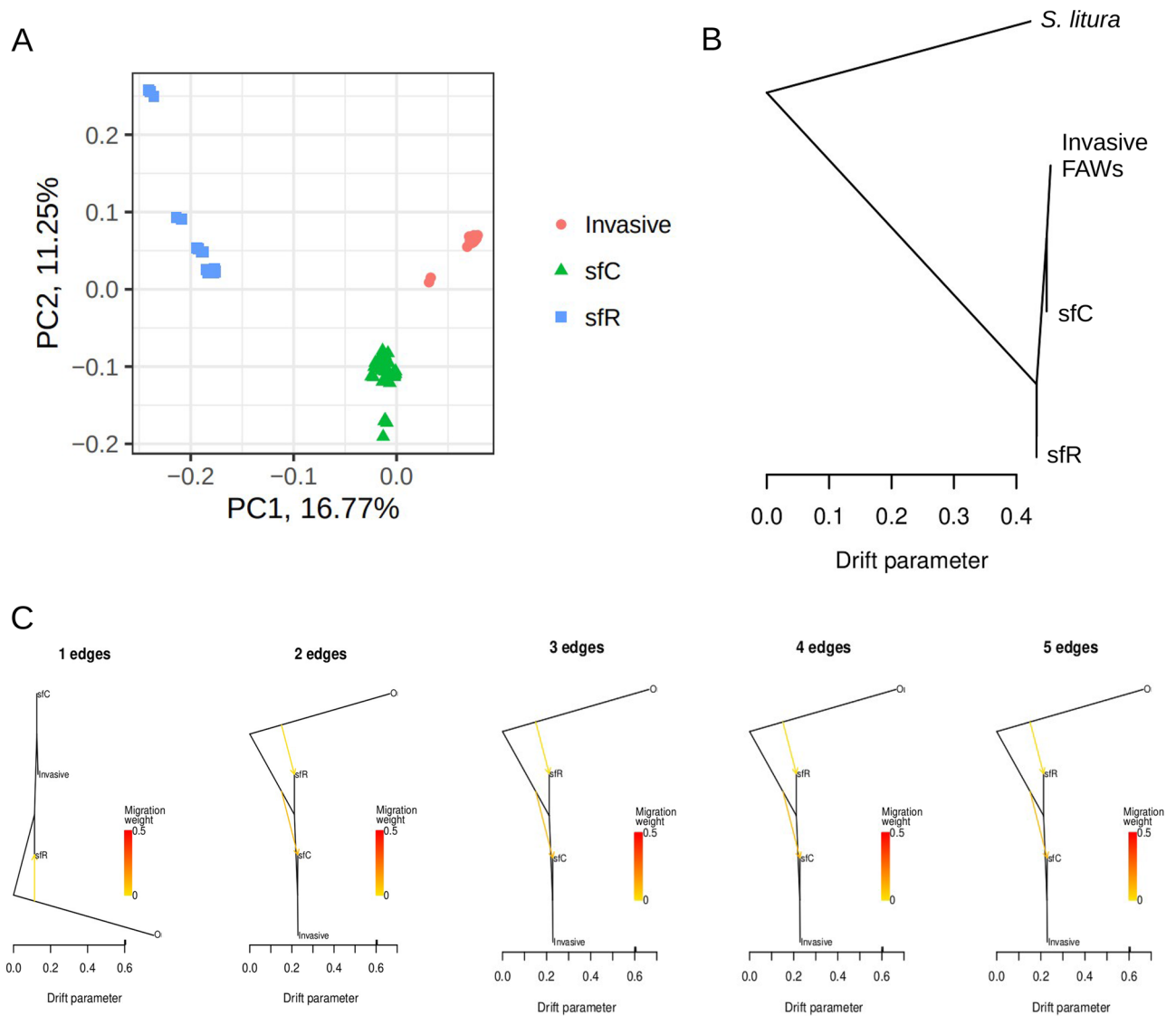
While FAWs exhibit extreme polyphagy by eating more than 353 plants in 76 families<sup>10</sup>, FAW consists of two strains with distinct ranges of host plants<sup>11,12</sup>. The corn strain (sfC) has a preference for corn, cotton, and sorghum, while the rice strain (sfR) prefers alfalfa, millet, pasture grasses, and rice<sup>13</sup>. Both sfC and sfR are sympatrically observed throughout almost the entire native habitat range. Reciprocal transplant experiments have demonstrated differential adaptation to host plants between sfC and sfR<sup>14</sup>. sfC and sfR showed a clear pattern of genetic differentiation across the whole genome sequences, indicating the presence of incipient speciation driven by differential host plant adaptation<sup>15,16</sup>. Since sfC and sfR are morphologically indistinguishable, molecular markers have been utilized to identify the strains. These molecular markers include the Z chromosome *triosephosphate isomerase* (TPI) gene<sup>17</sup> and the mitochondrial *cytochrome c oxidase subunit 1* (COX1) gene<sup>18</sup>.

The taxonomic identification of invasive FAWs is one of the most basic pieces of information to monitor invasive FAWs and assess their negative impacts. However, it is surprisingly unclear to which strains invasive FAWs belong, even though a large number of related studies have been performed. Previous population genetics studies suggested that invasive FAW populations are hybrids between the sfC and sfR strains based on the presence of sfC-type Z chromosome TPI and sfR-type mt-COX1 genes in invasive samples<sup>19–22</sup>. Population genetics studies based only on mt-COX1 genes concluded that sfR invaded<sup>23,24</sup>. However, with this conclusion, it is hard to explain why invasive FAWs are predominantly detected from sfC-preferred crops<sup>25</sup>, such as maize. Furthermore, it remains unclear whether invasive FAWs have the potential to cause a reduction in rice production, even

<sup>1</sup>DGIMI, Univ Montpellier, INRAE, Montpellier, France. <sup>2</sup>These authors contributed equally: Karine Durand and Hyerin An. ✉email: ki-woong.nam@inrae.fr

when a serious infestation of rice fields has never been reported. This ambiguity may be attributed to the limited discriminative capacity of the mt-COX1 or TPI genes in identifying strains. Our whole genome analyses have revealed that, while the TPI gene can reliably distinguish between strains, the mt-COX1 gene can only differentiate between two sub-strains within sfC<sup>16</sup>. Therefore, it is imperative to reevaluate the widespread belief that invasive FAWs are hybrids or sfRs through whole genome analyses.

We conducted whole-genome sequencing analysis on 177 samples collected globally, aiming to infer the evolutionary history of FAW invasions<sup>26</sup>. We showed clear genomic differentiation between the sfC and sfR strains in native FAW populations. We also observed that invasive populations are genomically closest to native sfC samples. We conclude that invasive populations should be considered as sfC strains rather than hybrids (Fig. 1A). This conclusion is indeed in line with the observation that invasive FAWs are observed almost exclusively from host plants preferred by sfC<sup>25</sup>. Since the observed genomic similarity reflects the averaged pattern across the whole genomes, it remains still possible that a part of invasive genomes originated from sfR through ancient gene flow while the majority of genomic sequences have sfC-type sequences. In such a scenario, invasive FAWs could be considered to be partial hybrids. This study aims to determine the strain of invasive populations by testing the existence of sfR-type genomic loci in invasive populations using publicly available whole genome sequences generated by our previous study. If the invasive FAW genome contains a part of sfR-derived sequences,



**Figure 1.** The phylogeny of FAW populations. **(A)** Principal component analysis revealed clear groupings of FAW individuals into three groups: sfC, sfR, and invasive populations. This figure was regenerated with the modification of Yainna et al.<sup>26</sup> to visualize the population structure of FAW. A clear grouping among sfC, sfR, and invasive populations is shown. Along the first principal component, sfC appeared closer to invasive populations than sfR. **(B)** The TreeMIX analysis indicated that ancestral FAWs split into sfR and sfC + invasive FAWs, suggesting that invasive FAWs originated from sfC. **(C)** Gene flow from sfR to invasive populations was not detected with a range of edge numbers.

infestation to rice could be considered for monitoring. However, if such a genetic introgression from sFR to the invasive populations is not observed, the risk of infesting rice should be considered to be low, as native sFCs<sup>16</sup>.

## Results

A phylogenetic tree inferred using TreeMix analysis showed that FAWs were divided into sFC + invasive populations and sFR (Fig. 1B). Together with Fig. 1A, these results showed that invasive populations are genetically closer to sFC than sFR, and invasive FAWs originated from sFC as shown by Yainna et al.<sup>26</sup>. According to the TreeMix analyses, gene flow from sFR to invasive populations was not detected in a range of edge numbers (Fig. 1C). The  $f_3$  statistic (sFC, sFR; invasive FAWs) was significantly higher than zero (0.038494;  $z$ -score = 32.544;  $p$ -value < 0.001), indicating that invasive FAWs are not genetically admixed between sFC and sFR.

Then, we tested the presence of invasive genomic loci that are more similar to sFR than sFC. The genomic average  $F_{ST}$  between invasive FAWs and sFR was 0.115, which was higher than the  $F_{ST}$  between invasive FAWs and sFC (0.0426).  $F_{ST}$  was also calculated from non-overlapping 500 kb windows across the genome to identify loci at which invasive populations have higher  $F_{ST}$  with sFC than with sFR. Among a total of 782 non-overlapping 500 kb windows across the genome (Fig. 2), only five windows showed such a pattern, corresponding to 0.639% of the whole genome sequences. The average  $F_{ST}$  across these five windows was 0.0532 and 0.0478 between sFC and invasive populations and between sFR and invasive populations, respectively. However, statistical differences between these two  $F_{ST}$  values were not supported ( $p$ -value = 0.249; one-sided randomization test with 1,000 replications). This result implies that no genomic locus in invasive populations has more similar sequences to sFR than sFC with statistical significance.

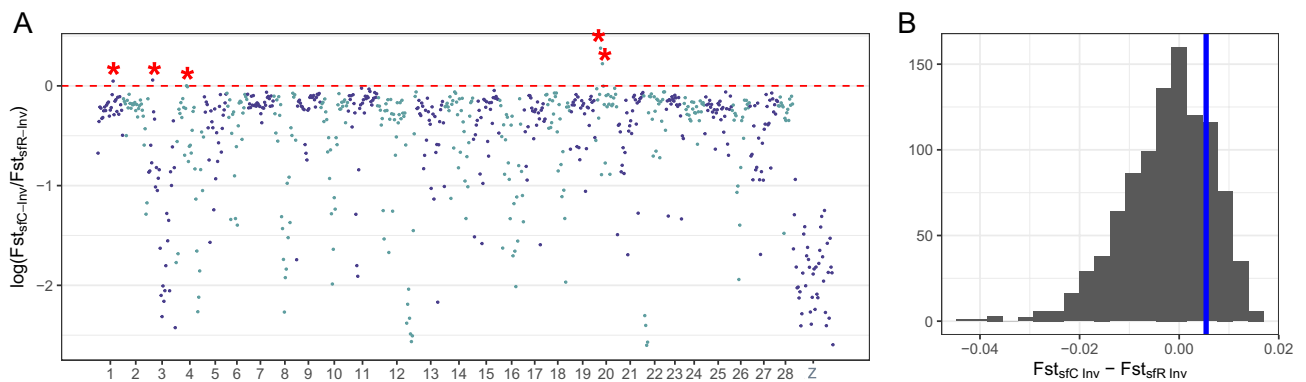
We also tested the existence of loci with genetic footprints of gene flow from sFR to the invasive population using  $D$  statistics. If incomplete lineage sorting caused an incongruent pattern between the phylogenetic tree shown in Fig. 1B and the gene tree inferred from the distribution of alleles, then, the same frequency of genetic variations will be observed between ABBA and BABA patterns, generating  $D$  statistics equal to zero<sup>27</sup> (Fig. 3A). Alternatively, if gene flows from sFR to invasive population existed, then, the ABBA pattern will outnumber the BABA pattern, making  $D$  statistics higher than zero.  $D$  statistics calculated from the whole genome sequences had a negative sign (-0.0540), implying that gene flow was not detected globally.

When  $D$  statistics were calculated from 10,000 SNV windows, only three loci in three chromosomes had  $D$  statistics higher than zero (Fig. 3B). These three loci windows correspond to 0.539% of all analyzed loci in the genomes. These loci do not have  $D$  statistics higher than zero with statistical significance ( $Z$ -test,  $p = 0.7225$ ).

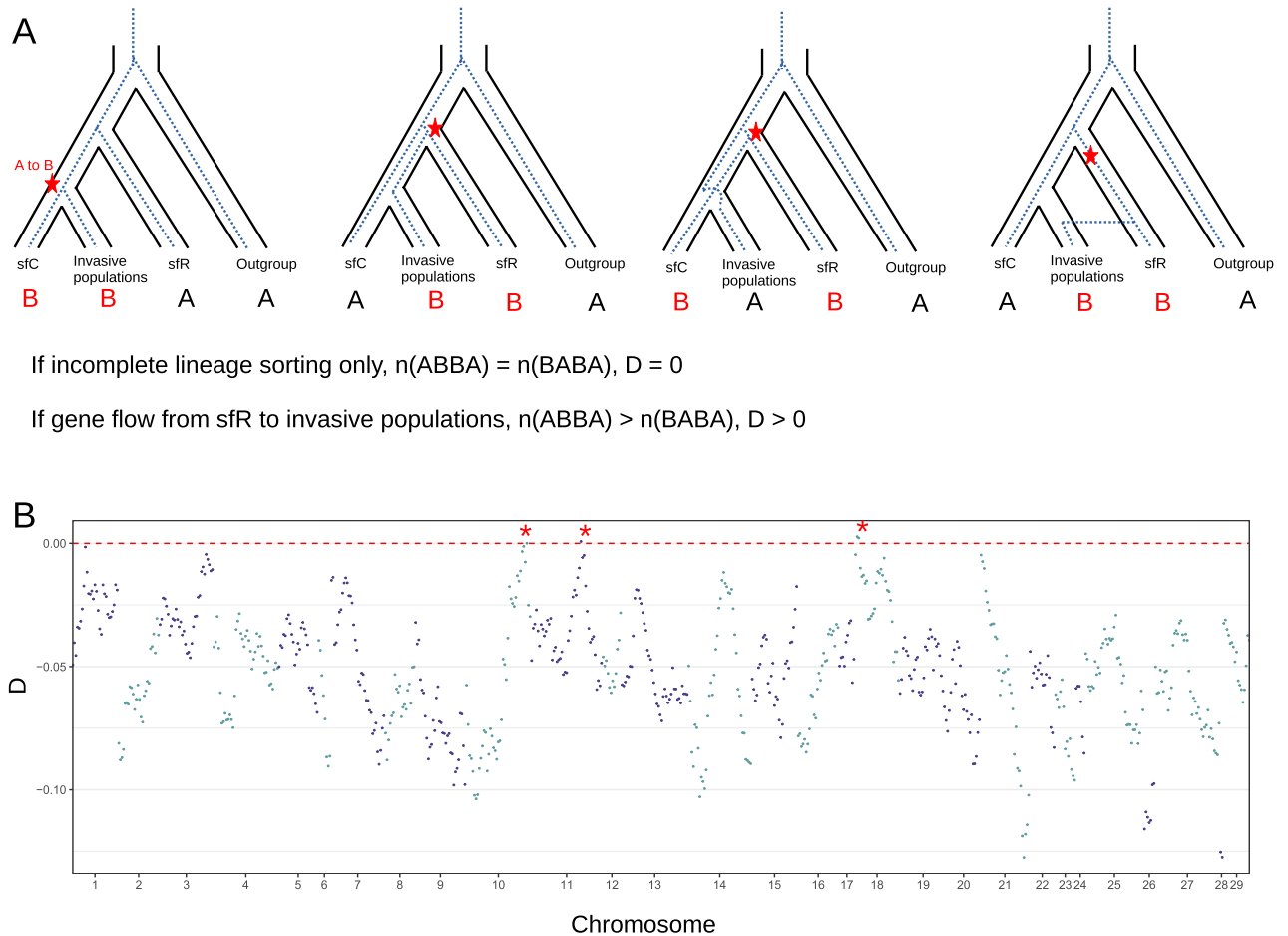
## Discussion

In this study, we aim to perform unambiguous taxonomic identification of invasive FAWs among sFC, sFR, and their hybrids using whole genome sequences from globally sampled 116 individuals. Phylogenetic analysis showed that invasive FAWs originated from sFC (Fig. 1), as shown by Yainna et al.<sup>26</sup>.  $F_{ST}$  statistics showed that invasive FAWs have no genomic locus that is closer to sFR than sFC (Fig. 2). According to the  $D$  statistics, gene flow from sFR to the invasive populations was not supported by any genomic locus (Fig. 3). Taken together, these results demonstrate that invasive FAWs correspond to sFC without genetic introgression from sFR. Therefore, we argue that invasive FAWs probably have a very limited capacity to infest rice, while this risk has been of great interest in invaded area<sup>28,29</sup>.

We argue that the misinterpretation of sFR-type mt-COX1 genes may have led to the conclusion that invasive FAWs are hybrids. This conclusion was based on the observation that the majority of invasive FAWs exhibit the sFC type according to the TPI marker and the sFR type according to the mt-COX1 genes<sup>19–22</sup>. Fiteni et al. showed through population genomics analyses conducted on FAW larvae collected from different host plants



**Figure 2.** Invasive populations are genetically closer to sFC than sFR across the whole genome. **(A)** The log-transformed ratio of  $F_{ST}$  between sFC and invasive populations ( $F_{ST_{sFC-inv}}$ ) to  $F_{ST}$  between sFR and the invasive population ( $F_{ST_{sFR-inv}}$ ) is represented. If  $F_{ST_{sFC-inv}} < F_{ST_{sFR-inv}}$  the log-transformed ratio will be below zero, indicated by the horizontal red dotted bar. Five 500 kb windows with  $F_{ST_{sFC-inv}} > F_{ST_{sFR-inv}}$  are marked with red asterisks. **(B)** The vertical blue bar represents the average difference between  $F_{ST_{sFC-inv}}$  and  $F_{ST_{sFR-inv}}$  across the five windows. The histogram illustrates the difference of  $F_{ST}$  calculated from two randomly generated native groups. The proportion of random groups with higher values than  $F_{ST_{sFC-inv}} - F_{ST_{sFR-inv}}$  corresponds to the  $p$ -values.



**Figure 3.** Testing gene flow from sfR to invasive populations. (A) If a mutation that changes from A to B originated from the common ancestor between sfC and invasive populations (indicated by the red stars), the genotype combination among sfC, invasive populations, sfR, and the outgroup will be BBAA. In the presence of incomplete lineage sorting, the gene tree may differ from the FAW strain tree in Fig. 1B. In such cases, when the frequency of ABBA and BABA will be the same. Then, D statistics, normalized differences between the frequency of ABBA and BABA will be zero. In the presence of gene flow from sfR to invasive populations, the invasive populations will acquire mutations generated in the sfR genealogy. As a result, the frequency of ABBA will be higher than that of BABA. Consequently, the D statistic will be higher than zero. (B) The majority of D statistics calculated across the genome are below zero, indicating a lack of gene flow. However, there are three loci with slightly higher D statistics than zero, as denoted by the red asterisks.

that FAW can be divided into two strains: sfC and sfR. Additionally, within the sfC subgroup, two sub-strains, mtA and mtB, have been identified, each with distinct mitochondrial and nuclear genomes<sup>16</sup>. Furthermore, they demonstrated that mitochondrial sequences are indistinguishable between mtB and sfR, whereas mtA and mtB exhibit clear differentiation in mitochondrial sequences. Therefore, the presence of sfR-type mt-COX1 genes in invasive FAWs may correspond to the mtB sub-strain, rather than sfR. Taking into consideration the absence of gene flow from sfR to invasive populations, as observed in this study, we argue that this possibility is likely true. Thus, we also raise the possibility that population genetics studies, by relying solely on mt-COX1 genes, might have been misled into concluding that sfR were invading<sup>23,24</sup>.

One possible criticism against our conclusion is the potential occurrence of gene flow from sfR to invasive populations that were not included in this study. In this case, a proportion of invasive populations might be hybrids. Zhang et al.<sup>30</sup> performed population genomics analysis using an independent resequencing dataset consisting of 280 FAW samples collected from five native areas (the mainland USA, Brazil, Guadeloupe, Argentina, and Puerto Rico) and eight invasive areas (China, Ethiopia, Ghana, Kenya, Malawi, Rwanda, South Africa, and Zambia). They reported that invasive populations are genetically closest to sfC while also reporting genomic differentiation among invasive populations, as shown by Yainna et al. The consistently observed patterns from different resequencing datasets covering wide geographical ranges suggest that a ghost population of sfR is unlikely to exist in the invaded area. Therefore, the inclusion of additional invasive populations is unlikely to change our conclusion in this study, even though we cannot absolutely exclude the possibility of ancient gene flow from sfR to unknown invasive populations.

In this study, we showed that invasive FAWs are pure sfC without gene flow from sfR. This result implies that invasive FAWs are not likely to cause massive infestation on sfR-preferred plants, such as rice or grasses. Even

though invasive FAWs may be anecdotally observed from rice, we do not believe that invasive FAWs will establish permanent populations on these plants without a special evolutionary force expanding the range of host plants. This argument does not imply that it is unnecessary to perform surveillance on sFR from invaded areas because sFR might arrive at invaded areas through additional introductions<sup>22</sup>. Instead, our study indicates that we need to consider the limited capacity of FAWs to infest rice when developing FAW monitoring or prevention plans.

## Methods

This study is based on the whole genome resequencing data generated by Yainna et al.<sup>26</sup>. The dataset consisted of 177 samples collected from 12 geographic locations, including Brazil, Florida, French Guinea, Guadeloupe, Mexico, Mississippi, and Puerto Rico for native areas and Benin, China, India, Malawi, and Uganda for invaded areas. The accession numbers to the raw reads were PRJNA494340, PRJNA577869, PRJNA639295, and PRJNA639296 in NCBI SRA. Please see Yainna et al. for detailed information. We excluded samples from Brazil, Malawi, and Uganda because the data was not publicly available when we performed the analyses<sup>8,31</sup>. Mexican samples were also excluded as they were phylogenetically distinct from the others. Additionally, one hybrid sample from Florida (FGJ4) and one sFR sample from a corn field in Mississippi (MS\_R\_R6) were excluded to avoid interference from local hybridization or migration between crop fields. The total number of analyzed samples was 116. These samples include 44 native sFC ones from Florida (13 samples), Mississippi (16), and Puerto Rico (15), 17 native sFR ones from Florida (10), French Guiana (3), and Guadeloupe (4), and 55 invasive ones from Benin (39), China (2), and India (14). The resequencing dataset includes 27,117,672 single nucleotide variations (SNVs) across whole-genome sequences from these 116 samples.

Principal component analysis was conducted using Plink v1.9<sup>32</sup>. A phylogenetic tree among sFC, sFR, and invasive populations was inferred using TreeMIX v1.13<sup>33</sup>. Whole genome sequences of *S. litura* were used as an outgroup (NCBI Accession: SRR5132437).  $F_{ST}$ <sup>34</sup> was calculated between sFC and invasive populations and between sFR and invasive populations from non-overlapping 500 kb windows using VCFtools v0.1.15<sup>35</sup>. The  $f_3$  statistic<sup>27</sup> was calculated to test the genetic admixture of invasive FAWs between sFC and sFR, using the following method. Initially, SNVs in the vcf file were excluded if the genotype was undetermined in any sample or if the SNV was non-biallelic using VCFtools v0.1.15<sup>35</sup>. Subsequently, the filtered VCF file was converted to PLINK format using VCFtools v0.1.15, followed by additional conversion to eigenstrat format using CONVERTF in the admixtools package v7.02<sup>27</sup>. Finally, the  $f_3$  statistic and the corresponding z-score were calculated using qp3Pop the admixtools package v7.02 in conjunction with admixr v0.9.1<sup>36</sup>. The p-value was derived from the Z table. The number of used SNVs was 232,518 in this analysis. D statistics indicating gene flow from sFR to invasive populations were computed from sliding windows using Dsuit v0.4 r38 with Dinvestigate option<sup>37</sup>. The number of SNVs was 10,000 for each window, and the step size was 1,000 SNVs. The statistical significance of D statistics greater than 0 was tested using Comp-D v6695c6b<sup>38</sup> from Z-statistics.

## Data availability

All data generated or analyzed during this study are included in this published article<sup>26</sup>. The NCBI SRA accession numbers of this project are PRJNA494340, PRJNA577869, PRJNA639295, and PRJNA639296. The computer programming scripts used during the current study are available from the corresponding author on request.

Received: 27 October 2023; Accepted: 5 March 2024

Published online: 08 March 2024

## References

- Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, 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 (2016).
- Day, R. et al. Fall armyworm: impacts and implications for Africa. *Outlooks Pest Manag.* **28**, 196–201 (2017).
- Nuss, E. T. & Tanumihardjo, S. A. Maize: A paramount staple crop in the context of global nutrition. *Compr. Rev. Food Sci. Food Saf.* **9**, 417–436 (2010).
- Bai-Zhong, Z. et al. Silencing of cytochrome P450 in *Spodoptera frugiperda* (Lepidoptera: Noctuidae) by RNA interference enhances susceptibility to chlorantraniliprole. *J. Insect Sci.* **20**, 12 (2020).
- Lv, S.-L. et al. Detection of ryanodine receptor target-site mutations in diamide insecticide-resistant *Spodoptera frugiperda* in China. *Insect Sci.* **28**, 639–648 (2021).
- Zhang, D. et al. Insecticide resistance monitoring for the invasive populations of fall armyworm, *Spodoptera frugiperda* in China. *J. Integr. Agric.* **20**, 783–791 (2021).
- Gui, F. et al. Genomic and transcriptomic analysis unveils population evolution and development of pesticide resistance in fall armyworm *Spodoptera frugiperda*. *Protein Cell* **1**, 1–19 (2020).
- Guan, F. et al. Whole-genome sequencing to detect mutations associated with resistance to insecticides and Bt proteins in *Spodoptera frugiperda*. *Insect Sci.* <https://doi.org/10.1111/1744-7917.12838> (2020).
- Zhao, Y.-X. et al. Susceptibility of fall armyworm, *Spodoptera frugiperda* (JE Smith), to eight insecticides in China, with special reference to lambda-cyhalothrin. *Pestic. Biochem. Physiol.* **168**, 104623 (2020).
- Montezano, D. G. et al. Host plants of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in the Americas. *Afr. Entomol.* **26**, 286–300 (2018).
- Pashley, D. P. Host-associated genetic differentiation in fall armyworm (Lepidoptera: Noctuidae): A sibling species complex?. *Ann. Entomol. Soc. Am.* **79**, 898–904 (1986).
- Pashley, D. P. & Martin, J. A. Reproductive incompatibility between host strains of the Fall Armyworm (Lepidoptera: Noctuidae). *Ann. Entomol. Soc. Am.* **80**, 731–733 (1987).
- Juárez, M. L. et al. Host association of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) corn and rice strains in Argentina, Brazil, and Paraguay. *J. Econ. Entomol.* **105**, 573–582 (2012).
- Orsucci, M. et al. Transcriptional differences between the two host strains of *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Peer Community J.* **2**, 1–10 (2022).

15. Durand, K., Yainna, S. & Nam, K. Incipient speciation between host-plant strains in the fall armyworm. *BMC Ecol. Evol.* **22**, 52 (2022).
16. Fiteni, E. *et al.* Host-plant adaptation as a driver of incipient speciation in the fall armyworm (*Spodoptera frugiperda*). *BMC Ecol. Evol.* **22**, 133 (2022).
17. Nagoshi, R. N. The fall armyworm *Triosephosphate Isomerase* (Tpi) gene as a marker of strain identity and interstrain mating. *Ann. Entomol. Soc. Am.* **103**, 283–292 (2010).
18. Pashley, D. P. Host-associated differentiation in armyworms (Lepidoptera: Noctuidae): An allozymic and mtDNA perspective. in *Electrophoretic Studies on Agricultural Pests*, vol. 39, 103–114 (Clarendon Press, 1989).
19. Nagoshi, R. N. *et al.* Fall armyworm migration across the Lesser Antilles and the potential for genetic exchanges between North and South American populations. *PLoS ONE* **12**, e0171743 (2017).
20. Nagoshi, R. N. *et al.* Analysis of strain distribution, migratory potential, and invasion history of fall armyworm populations in northern Sub-Saharan Africa. *Sci. Rep.* **8**, 3710 (2018).
21. Nagoshi, R. N., Goergen, G., Plessis, H. D., van den Berg, J. & Meagher, R. 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 (2019).
22. Nagoshi, R. N. *et al.* Genetic studies of fall armyworm indicate a new introduction into Africa and identify limits to its migratory behavior. *Sci. Rep.* **12**, 1941 (2022).
23. Herlinda, S., Suharjo, R., Elbi Sinaga, M., Fawwazi, F. & Suwandi, S. First report of occurrence of corn and rice strains of fall armyworm, *Spodoptera frugiperda* in South Sumatra, Indonesia and its damage in maize. *J. Saudi Soc. Agric. Sci.* **21**, 412–419 (2022).
24. Yousaf, S., Rehman, A., Masood, M., Ali, K. & Suleman, N. Occurrence and molecular identification of an invasive rice strain of fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae) from Sindh, Pakistan, using mitochondrial cytochrome c oxidase I gene sequences. *J. Plant Dis. Prot.* **129**, 71–78 (2022).
25. Nagoshi, R. N. *et al.* The fall armyworm strain associated with most rice, millet, and pasture infestations in the Western Hemisphere is rare or absent in Ghana and Togo. *PLoS ONE* **16**, e0253528 (2021).
26. Yainna, S. *et al.* The evolutionary process of invasion in the fall armyworm (*Spodoptera frugiperda*). *Sci. Rep.* **12**, 21063 (2022).
27. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
28. Ali, M. P. *et al.* Development and demographic parameters of fall armyworm (*Spodoptera frugiperda* J. E. Smith) when feeding on rice (*Oryza sativa*). *CABI Agric. Biosci.* **4**, 29 (2023).
29. He, L. *et al.* Larval diet affects development and reproduction of East Asian strain of the fall armyworm, *Spodoptera frugiperda*. *J. Integr. Agric.* **20**, 736–744 (2021).
30. Zhang, L. *et al.* Global genomic signature reveals the evolution of fall armyworm in the Eastern hemisphere. *Mol. Ecol.* **32**, 5463–5478 (2023).
31. Tay, W. T. *et al.* Global population genomic signature of *Spodoptera frugiperda* (fall armyworm) supports complex introduction events across the Old World. *Commun. Biol.* **5**, 1–15 (2022).
32. Renteria, M. E., Cortes, A. & Medland, S. E. Using PLINK for genome-wide association studies (GWAS) and data analysis. *Methods Mol. Biol.* **1019**, 193–213 (2013).
33. Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
34. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
35. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
36. Petr, M., Vernot, B. & Kelso, J. admixr: R package for reproducible analyses using ADMIXTOOLS. *Bioinformatics* **35**, 3194–3195 (2019).
37. Malinsky, M., Matschiner, M. & Svardal, H. Dsuite: Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* **21**, 584–595 (2021).
38. Musmann, S. M., Douglas, M. R., Bangs, M. R. & Douglas, M. E. Comp-D: A program for comprehensive computation of D-statistics and population summaries of reticulated evolution. *Conserv. Genet. Resour.* **12**, 263–267 (2020).

## Acknowledgements

The study is supported by a fund from the department of Santé des Plantes et Environnement at Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (Resistome).

## Author contributions

KD, HA, and KN performed the analyses. KN designed the study. KD generated Figs. 1 and 2. HA generated Fig. 3. KD, HA, and KN wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to K.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024