Rapid evolution of pesticide resistance via adaptation and interspecific introgression in a major North American crop pest

Key words: introgression, *Helicoverpa zea*, *Helicoverpa armigera*, selective sweep, isolation by distance, pesticide resistance, Bt resistance, population genomics, pests, invasive species

Henry L. North[*1†], Zhen Fu[*2,3], Richard Metz[4], Matt A. Stull[4], Charles D. Johnson[4], Xanthe Shirley[5], Kate Crumley[6], Dominic Reisig[7], David L. Kerns[2], Todd Gilligan[8], Tom Walsh[9], Chris D. Jiggins[1†], Gregory A. Sword[2†]

* These authors contributed equally to this work

[1]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
[2] Department of Entomology, Texas A&M University, College Station, TX 77843, USA
[3]Bioinformatics and Biostatistics Core, Van Andel Institute, Grand Rapids, MI, 49505, USA
[4]Texas A&M AgriLife Genomics and Bioinformatics Service, College Station, TX, USA, 77843
[5]USDA Animal and Plant Health Inspection Service, College Station, TX, USA
[6]Texas A&M Agrilife Extension, Wharton, TX, USA
[7]North Carolina State University, Entomology and Plant Pathology, Plymouth, NC, USA
[8]USDA Animal and Plant Health Inspection Service, Fort Collins, CO, USA
[9]Black Mountain Laboratories, Commonwealth Scientific and Industrial Research Organization, Canberra, Australia

[†] Corresponding authors (gasword@tamu.edu, cj107@cam.ac.uk, hln33@cam.ac.uk)

Abstract

Insect crop pests threaten global food security. This threat is amplified through the spread of non-native species and the evolution of pesticide resistance, which can be introduced to a population though *de novo* mutation or gene flow. We investigate these processes in an economically important noctuid crop pest, *Helicoverpa zea*, which has rapidly evolved resistance to several pesticides. Its sister species *Helicoverpa armigera*, first detected as an invasive species in Brazil in 2013, introduced the pyrethroid resistance gene *CYP337B3* to South American *H. zea* via introgression. To understand whether this contributes to pesticide resistance in North America, we sequenced 237 *H. zea* genomes across 10 sample sites in the US. First, we report *H. armigera* introgression into the North American *H. zea* population. Two individuals sampled in Texas in 2019 carry *H. armigera* haplotypes in a 4Mbp region containing *CYP337B3*. Second, we show that the remarkable dispersal ability of *H. zea* results in a panmictic North American population. Third, we detect signatures of selection in non-admixed *H. zea*, identifying a selective sweep at a second pesticide resistance locus with a similar name: *CYP333B3*. We estimate that its derived allele conferred a ~4.9% fitness advantage and show that this estimate explains independently observed rare nonsynonymous *CYP333B3* mutations approaching fixation over a ~20-year period. We also detect putative signatures of selection at a kinesin gene associated with Bt resistance. Our results show that pesticide resistance in *H. zea* evolved rapidly and recently via two independent mechanisms: interspecific introgression and rapid intraspecific adaptation.

Introduction

Insect pests destroy 5-20% of major grain crop production, and losses are set to increase substantially over coming decades as a result of climate change, the evolution of pesticide resistance, and the spread of invasive species via international trade routes (Deutsch et al. 2018; Gould et al. 2018; Paini et al. 2016). Understanding the role that evolution plays in the ecology of pests and their outbreaks is therefore a priority (Luke et al. 2023; Karlsson Green et al. 2020). The short generation times and large effective population sizes of many invertebrate pest species results in high rates of molecular evolution and rapid allele frequency shifts in response to natural selection (Thomas et al. 2010; Petit & Barbadilla 2009). At the same time, extreme selective regimes imposed by pesticide exposure can result in rapid adaptation (Hawkins et al. 2019). The strong dispersal ability of many insect pests, especially in the context of contiguous habitat in monoculture, means that pest populations often exist as highly connected metapopulations that cover agricultural landscapes (Mazzi & Dorn 2012). As a consequence, fitness-enhancing alleles can readily spread across space (McDonald & Linde 2003). Human activity can also mediate dispersal at larger geographic scales (*i.e.*, across agricultural systems and continents) such that insecticide resistance can rapidly arise through gene flow – a process significantly faster than adaptation from de-novo mutation (Tay & Gordon 2019). The extent of global trade networks means that many closely related pest species come into secondary contact, opening up the possibility of adaptive introgression not only between populations, but also between divergent ecotypes and species (Valencia-Montoya et al. 2020; Song et al. 2011). Together these factors enhance the evolutionary potential of insect pests, with two key consequences. First, adaptive responses can occur at timescales relevant to year-to-year pest management strategies. Second, the global connectedness of many pest populations means that such strategies must be multilateral. Large-scale genomic monitoring is widely accepted as a promising emerging means of informing management action to address these consequences.

Population genetics has long been used as a tool for quantifying evolutionary change in agricultural pests, especially with respect to insecticide resistance (Mallet 1989). The additional information and contiguity of resolution offered by genome-resequencing data have created renewed interest in this field, and genomic approaches are clearly emerging as a key tool for monitoring pest populations under both proactive and reactive management plans (Sherpa & Després 2021; Hamelin & Roe 2020; Neafsey et al. 2021; North et al. 2021). Recent studies have demonstrated the use of population genomics approaches to define management units by quantifying population connectivity (Chen et al. 2021; Paris et al. 2022), identifying loci associated with the evolution of pesticide resistance by inferring the action of selection (Love et al. 2023), and reconstructing the spatial spread of species or alleles of interest (Tay et al. 2022). With appropriate analysis and sampling design, population genomics can be used to extract otherwise-inaccessible biological information to understand the evolutionary history of pest populations and inform management plans.

Helicoverpa zea, commonly known as the corn earworm, is a polyphagous noctuid moth common throughout the Americas. A notorious pest of maize and cotton, *H. zea* is one of the most economically significant crop pests in the agricultural powerhouses of Brazil and the United States (Olivi et al. 2019; Fitt 1989; Cook & Threet 2019; Olmstead et al.; Cunningham & Zalucki 2014). Although maize is its dominant host plant, larvae are known to feed on at least 122 other species, of which 29 are major crops including wheat, soy, rice, sorghum, and tomato (Cunningham & Zalucki 2014). Larvae tend to feed on the fruiting body of the plant, thereby directly damaging produce (Luttrell & Jackson). Generation times vary depending on latitude (5-10 generations/yr at lower altitudes), though facultative diapause enables pupa to persist underground up to at least 40°N during winter (Hardwick 1965; Morey 2010; Parajulee et al. 2004). Adults are highly effective long-distance dispersers, expanding northward into extensive areas of maize production during summer to a latitude of ~52°N in flights large enough to detect using ground-based radar (Jones et al. 2018;

Westbrook 2008). The species' range is expected to expand twofold by 2099 as warmer winters reduce the number of lethal low-temperature events (Lawton et al. 2022). Repeated admixture due to seasonal re-establishment from southern populations, combined with long-range dispersal over highly connected agricultural habitat, result in a highly connected and genetically diverse metapopulation in the north (Seymour et al. 2016; Margosian et al. 2009).

Multiple studies have found that the local distribution of crops producing *Bacillus thuringiensis* (Bt) toxins in a given year can predict *H. zea* damage in subsequent years (Arends et al. 2021, 2022). This observation implies that selection results in geographically localised phenotypic change between generations, so parent-offspring dispersal should primarily occur at the same geographic scale. In contrast, population genetics studies have reported effective panmixia across the North American range, though there is mixed evidence for this observation, and the only study to employ whole-genome data compared just two sample sites (Perera et al. 2020; Seymour et al. 2016; Taylor et al. 2021). Increased sampling effort — in terms of both geographic range and number of loci — can reveal population structure that is otherwise undetectable, as demonstrated in studies of *Helicoverpa armigera* (Jin et al. 2023; Zhang et al. 2022). Characterising the landscape of effective migration is key to understanding how rapidly adaptive variants underlying pesticide resistance can spread.

H. zea has evolved resistance to several pesticides. In the United States, the organochloride DDT was effective for *H. zea* control from its post-war implementation until the 1960s; resistance to methyl parathion (an organophosphate) introduced 1960s was initially detected in the 1970s; several pyrethroids introduced in the 1970s were effective until resistance started to become apparent in the 1990s and 2000s (Abd-Elghafar et al. 1993; Walsh et al. 2022). The specific mode of action of these pesticides means that resistance phenotypes could in many cases be underlain by mutations at one or few loci and therefore evolve rapidly in response to strong selection (Ibrahim et al. 2015). As a member of the 'megapest' genus of polyphagous herbivores, *Helicoverpa*, *H. zea* may be particularly well-equipped to evolve resistance due to its ecology and its capacity to metabolise a wide array of host plant defences (Gordon et al. 2010; Mallet 1989; Bras et al. 2022). Although many phenotypes with different genetic architectures can underly resistance to a given pesticide, certain gene families have been repeatedly implicated as targets of selection. This is true of cytochrome P450 genes involved in xenobiotic metabolism not only in *Helicoverpa* but in pest species spanning the tree of life (Kreiner et al. 2019; Nauen et al. 2022).

By the 1990s it became clear that the evolution of resistance was outpacing the development of novel pesticides, highlighting the need for a strategic shift toward integrated pest management. New approaches implemented for lepidopteran pests included the use of cotton and maize crops engineered to produce Bt toxins, which today constitute 82% of US maize crops and 88% of cotton (Reisig et al. 2022). In contrast to the nerve- and muscle-targeting insecticides discussed above, Bt toxins induce pore formation in the midgut membrane. Resistance to these toxins therefore requires selection on an entirely different set of loci. The most successful management plans made use of Bt crops expressing multiple toxins (*i.e.*, different Cry or Vip proteins) at high concentrations (>25x the dose required to kill susceptible insects) planted among non-Bt refuges in which rare resistant individuals could reproduce with susceptible mates. For insect pests generally, this approach was broadly successful at minimising unidirectional selection pressures, reducing net pesticide use and slowing the rate of resistance evolution. However, *H. zea* is among a few pest species to have evolved Bt resistance in the field. Cry1Ac resistance was reported in the early 2000s, at which point multi-toxin crops additionally expressing Cry2 were planted (Ali et al. 2006). By 2016, resistance to Cry1 toxins had become common throughout the US and Cry2 resistance was emerging (Dively et al. 2016; Reisig et al. 2018). Recent studies concluded that resistance to Cry2Ab2 has become common, and that selection was ongoing as of 2019 (Yu et al. 2021; Huang et al. 2023). The genetic basis of Bt

resistance is known to be complex and likely arose from standing variation, with unique genetic architectures underlying resistance to different Cry toxins (Taylor et al. 2021; Benowitz et al. 2022).

In addition to intraspecific adaptation, a major concern for the spread of pesticide resistance in North American *H. zea* is through interspecific introgression from its sister species *H. armigera*. Commonly known as the cotton bollworm, *H. armigera* has a broad Afro-Eurasian native range. Among the most economically damaging crop pests in the world, *H. armigera* is more polyphagous and resistant to a substantially broader array of pesticides compared with *H. zea* (Cunningham & Zalucki 2014). *H. armigera* was first detected in Brazil in 2013, where the two species hybridized (Tay et al. 2013; Anderson et al. 2018; Ivey & Hillier 2023). This resulted in the adaptive introgression of the *CYP337B3* gene into South American *H. zea* populations (Valencia-Montoya et al. 2020). *CYP337B3*, otherwise absent in *H. zea*, is the result of unequal crossover from two other cytochrome P450 genes and confers fenvalerate resistance (Joußen et al. 2012). The variant has arisen multiple times in *H. armigera*, and can encode resistance to various pyrethroids including cypermethrin and deltamethrin (Durigan et al. 2017; Rasool et al. 2014)

H. armigera is now established throughout much of South and Central America. The risk of *H. armigera*, or admixed *H. armigera-zea* individuals, spreading into suitable North American habitat is considered high, and such an event would put at risk US crop production valued at $USD 78 billion per annum (Kriticos et al. 2015). *H. armigera* has been intercepted at US ports more than 1000 times, suggesting that introduction via shipping routes from its broad Afro-Eurasian range is also a substantial risk (Kriticos et al. 2015). Since the two species are difficult to distinguish phenotypically, and because of the extent of interspecific admixture in South America, the detection of invasive *H. armigera* alleles into North requires genetic surveillance of native *H. zea* populations. To date, there have been no published reports in the scientific literature of *H. armigera* establishing on the North American mainland. In 2015, three specimens captured in Florida carried the *H. armigera* COI haplotype (Tembrock et al. 2019), and multiple adults were captured in areas adjacent to Chicago O'Hare International Airport (USDA APHIS), though these are isolated incidents. There have been unpublished reports of *CYP337B3* detected in *H. zea* survey samples, though this observation may result from parallel evolution, as observed in *H. armigera* (Rasool et al. 2014).

Here, we use a population genomics approach to (*i*) test for signatures of introgression from *H. armigera* into North American *H. zea*, (*ii*) characterise effective migration across space to understand how rapidly resistance alleles may spread, and (*iii*) conduct a genome-wide scan for evidence of selective sweeps at known pesticide and Bt resistance loci. To achieve this, we resequenced the genomes of 237 *H. zea* individuals across 10 locations in the US collected in 2019 (see Methods).

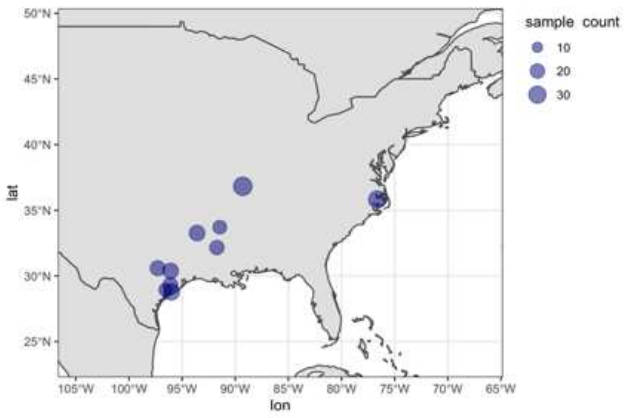
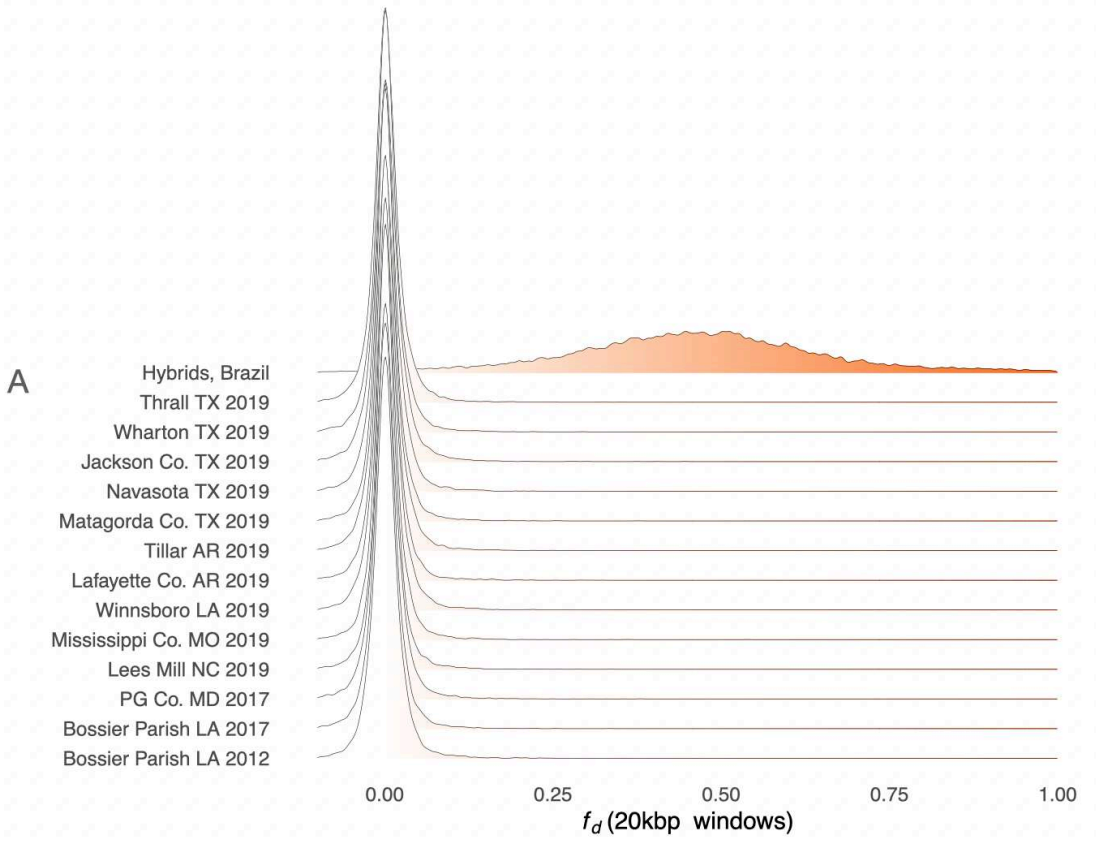


Figure 1: **Individual sampling effort at 10 sites in 2019.** Sample site information is detailed in Supplementary Table S1.

Results

Evidence for introgression of pesticide resistance genes into North American H. zea

We tested for the presence of *H. armigera* ancestry by calculating $\widehat{f_d}$ in 20kbp windows for individuals from each sample, using *H. zea* samples collected in Louisiana in 2002 as representative non-admixed samples and *H. punctigera* as an outgroup. We repeated this test for hybrids sampled in Brazil used as a positive control. The distribution of $\widehat{f_d}$ was centred on zero for all test sets apart from the positive control (Figure 2A). The same result is seen in samples reported in Taylor *et al.* (2021), collected in the US in 2012 and 2017. These results indicate balanced proportions of ABBA and BABA patterns, indicating little to no introgression at any sample site, at least compared to the magnitude of introgression seen in Brazil. However, $\widehat{f_d}$ is elevated on chromosome 15 among samples collected in 2019 in Jackson County, TX (Figure 2B). Computing $\widehat{f_d}$ for each of the 16 samples from this site individually shows that the pattern is driven by extreme values of $\widehat{f_d}$ in only two individuals (Supplementary Figure S1).



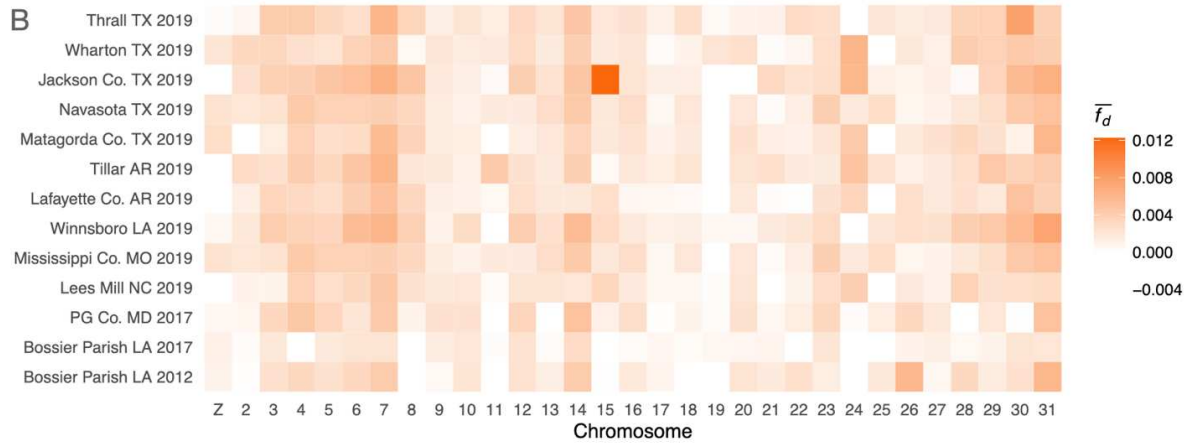


Figure 2: **No evidence of *H. armigera* in *H. zea* except for on chromosome 15 among a subset of *H. zea* individuals collected in Jackson County, TX, in 2019.** A: Distribution of $\widehat{f_d}$ calculated in 20kbp windows where P1: *H. zea* sampled in 2002, P3: *H. armigera*, outgroup: *H. punctigera*. The statistic was calculated for 14 different P2 sets: 10 sets of *H. zea* samples collected at different sites in 2019, 2 sets of *H. zea* collected in 2017, one set of *H. zea* collected in 2012, and a positive control of 9 individuals sampled in Brazil shown to be admixed individuals carrying both *H. armigera* and *H. zea* ancestry. B: Data presented as the mean per chromosome.

For those individuals, $\widehat{f_d}$ peaks at the *CYP337B3* locus (Figure 3). Nucleotide variation for these individuals is reduced across most of Chromosome 15 relative to *H. armigera*, but matches *H. armigera* in a terminal ~4Mbp region around *CYP337B3*. Given that the effective population size of *H. armigera* is twice that of *H. zea* (Anderson et al. 2018), this pattern suggests *H. armigera* ancestry dominates in this region. Based on patterns $\widehat{f_d}$ and π, we defined two segments of chromosome 15: A (0-9Mbp), which shows no signs of admixture, and B (>9Mbp) in which *H. armigera* haplotypes have introgressed.

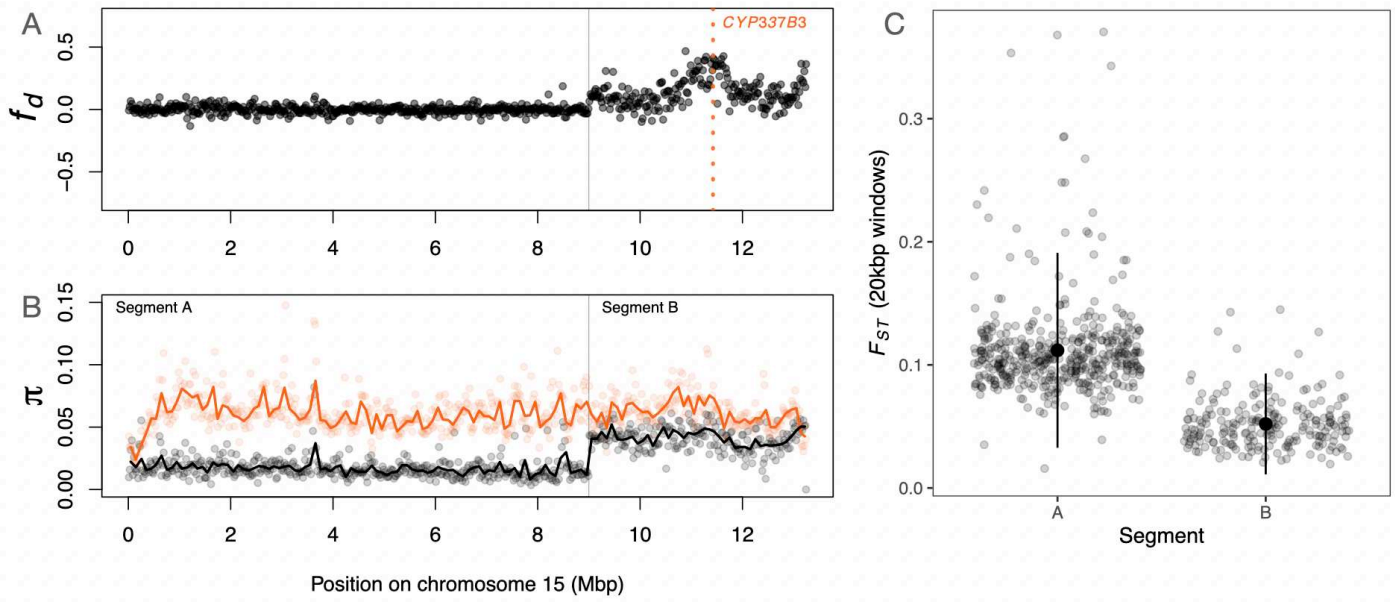


***Figure 3*: Introgression of *CPY337B3*.** A: $\widehat{f_d}$ calculated in 20kbp windows along chromosome 15, where P1 is *H. zea* sampled in 2002, P2 are Ja15 and Ja25 in supplementary Figure S1, P3 is *H. armigera*, and

the outgroup is *H. punctigera*. **B**: Nucleotide diversity (π) calculated in 20kbp windows (points) and 100kbp windows (lines) for the admixed individuals (Ja15 and Ja23; black) and for *H. armigera* (orange). **C**: Genetic differentiation (F_{ST}), calculated in 20kbp windows, between the admixed individuals and *H. armigera* in the chromosomal segments labelled in (B).

We reasoned that if segment B consists of largely *H. armigera* ancestry, genetic differentiation and divergence from *H. armigera* should be lower in this region. Segment B shows reduced genetic differentiation relative to *H. armigera* ($\overline{F_{ST}}$= 0.11 compared to 0.05 in Segment A; p<0.01, t = 25.606, df = 656, Welsch two-sample T test; Figure 3C). Genetic divergence also differed by segment ($\overline{d_{xy}}$= 0.086 and 0.073 respectively; p<0.01, t = 10.495, df = 515.7; Supplementary Figure S2); this this difference was less pronounced, as expected given that differentiation between species should accumulate slower than differentiation.

These differences were apparent when visualising the data with principal components analysis (PCA). At segment A, the two admixed individuals cluster completely within other *H. zea* samples from 2019 in principal component space, whereas at segment B the individuals are closer to known Brazilian hybrid samples (Figure 4).

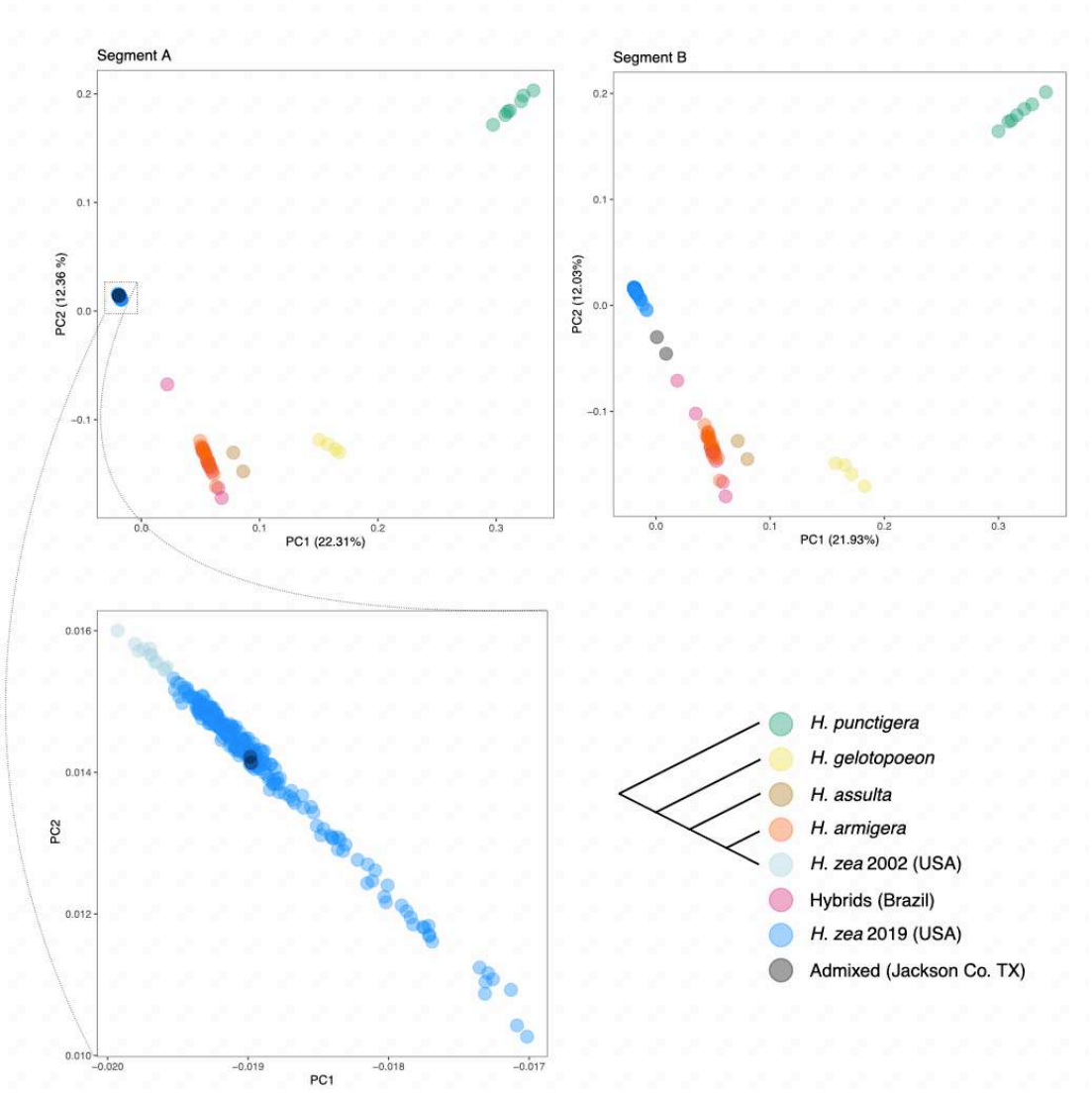


Figure 4: Visualising local ancestry with Principal Components Analysis. Principal components 1 and 2 calculated using SNPs form segments A and B of chromosome 15 (see Figure 3). Phylogeny in bottom right panel indicates the consensus species tree. In segment A, the two admixed samples (black points) cluster with other *H. zea* samples collected in 2019. In Segment B, the two admixed samples move in PC space toward *H. armigera* samples and admixed samples from Brazil.

We next sought to determine whether the admixed samples carried the *H. armigera CYP337B3* gene, and if so, to use the gene tree to identify a potential *H. armigera* source population. This is because *CYP337B3* has arisen independently in multiple *H. armigera* populations. To investigate this, we reconstructed a maximum likelihood gene tree at the *CYP337B3* locus (HaChr13:11436565-11440168, as mapped by Anderson *et al.* (Anderson et al. 2018) and used by Valencia Montoya *et al.* (Valencia-Montoya et al. 2020)), comparing the admixed samples to publicly available *H. armigera* samples representing the breadth of the species' phylogeographic diversity. The two admixed individuals form a clade with *H. armigera* samples at the *CYP337B3* locus (Figure 5). In 100% of bootstrap iterations, non-admixed *H. zea* samples were split from the clade consisting of admixed individuals and *H. armigera* samples. However, bootstrap support values were low within the *H. armigera* clade; there was no resolution to distinguish source populations. This result is consistent with low population structure within *H. armigera* and directional selection at this locus across the species' native and invasive range (Ni et al. 2023; Jin et al. 2023).

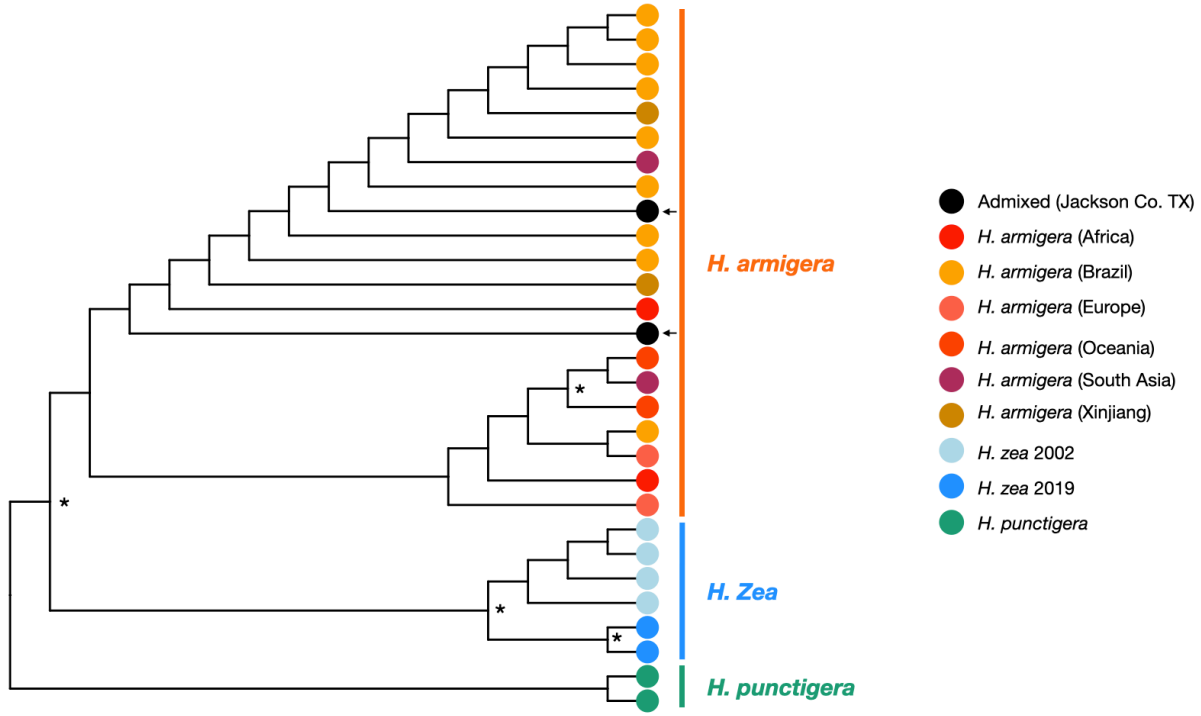


Figure 5: At the CYP337B3* locus, *admixed samples form a clade with H. armigera samples. Maximum likelihood cladogram rooted with *H. punctigera* samples. Asterixis denote nodes with >95% bootstrap support. Arrows indicate admixed samples. Horizontal lines group species-level clade. There was consistent bootstrap support separating species-level clades and *H. zea* samples from 2002 versus 2019, but not within *H. armigera*.

Together, these results suggest that two of 237 individuals sampled are admixed, that introgression in these two samples is concentrated around the *CYP337B3* locus, and that both samples carry at least one *H. armigera CYP337B3* allele. Through its semidominant mode of action both individuals are therefore resistant to fenvalerate (Joußen et al. 2012).

High connectivity across the North American *H. zea* metapopulation

We next sought to characterise the extent of allele sharing among non-admixed samples across the landscape to better understand how rapidly adaptive alleles (including those introduced from *H. armigera*) could spread. There was no positive relationship between geographic distance and genetic differentiation between sampling sites (one-sided Pearson's product moment correlation for genome-wide average F_{ST}: t = -2.8132, 43 degrees of freedom, p = >0.05), and the overall level of genetic differentiation was very low between sample sites (F_{ST}< 0.01), so the North American *H. zea* metapopulation shows no signs of isolation by distance (Figure 6). In effect, individuals collected at all sample sites have the properties of a single Wright-Fisher-like population, meaning that segregating sites among samples from all locations could be used to detect selective sweeps. This provided us with substantial statistical power for our third aim.

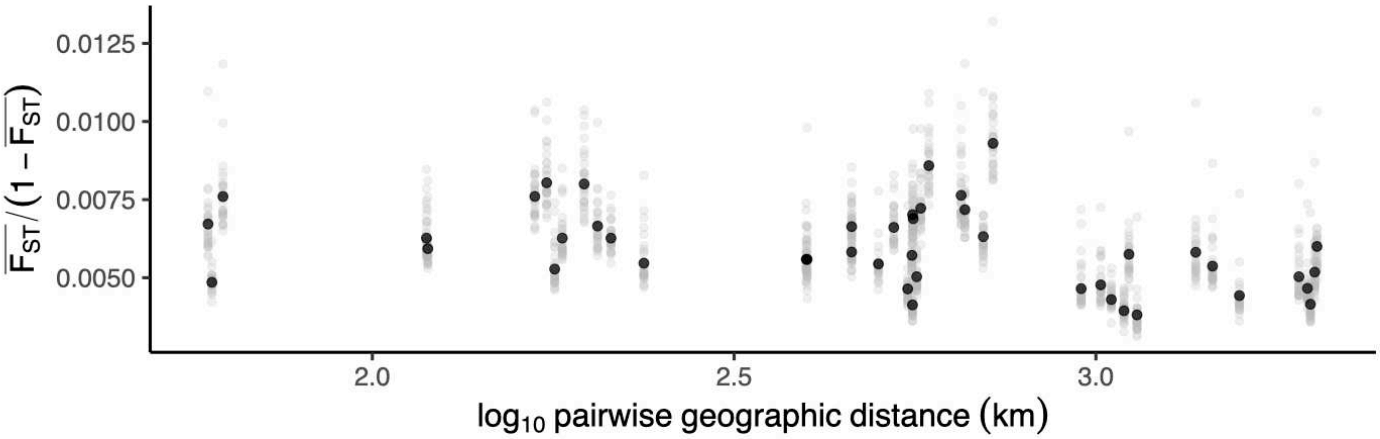


Figure 6: Low genetic differentiation irrespective of geographic distance among non-admixed H. zea. Pairwise genetic differentiation regressed on pairwise geographic distance for each of 45 possible comparisons among 10 sample sites. There is no positive relationship. Grey points are results for each of 31 chromosomes; black points are genome-wide averages. Genetic differentiation was calculated as mean F_{ST} across 20kbp windows.

Signatures of selection throughout the genome

Given the absence of population structure among our samples, we included all individuals in a genome-wide scan for selective sweeps. We identified putative sweeps as extreme outliers for the composite likelihood ratio (CLR) implemented in *SweepFinder2* (see Methods). Three sweeps occurred on chromosomes 13, 15 and 25 (Figures 7-8; Supplementary Figure S5). The sweep on chromosome 15 occurred within Segment A (Figure 3) and showed no signs of introgression in any samples.

Several genes within sweep regions have potential roles in immune response and pesticide resistance (Supplementary Table S4). The sweep on chromosome 15 (Figure 7) contained six unique gene annotations, with the vast majority of CLR outliers observed within one gene – *Hyd*-like – encoding an immune-associated E3 ubiquitin-protein ligase (Cammarata-Mouchtouris et al. 2020). The sweep on chromosome 25 contained 7 functional annotations (Figure 6) and the sweep on chromosome 13, contained two genes including a different cytochrome P450 gene, *CYP333B3* (Supplementary Table 3; Figure 8).

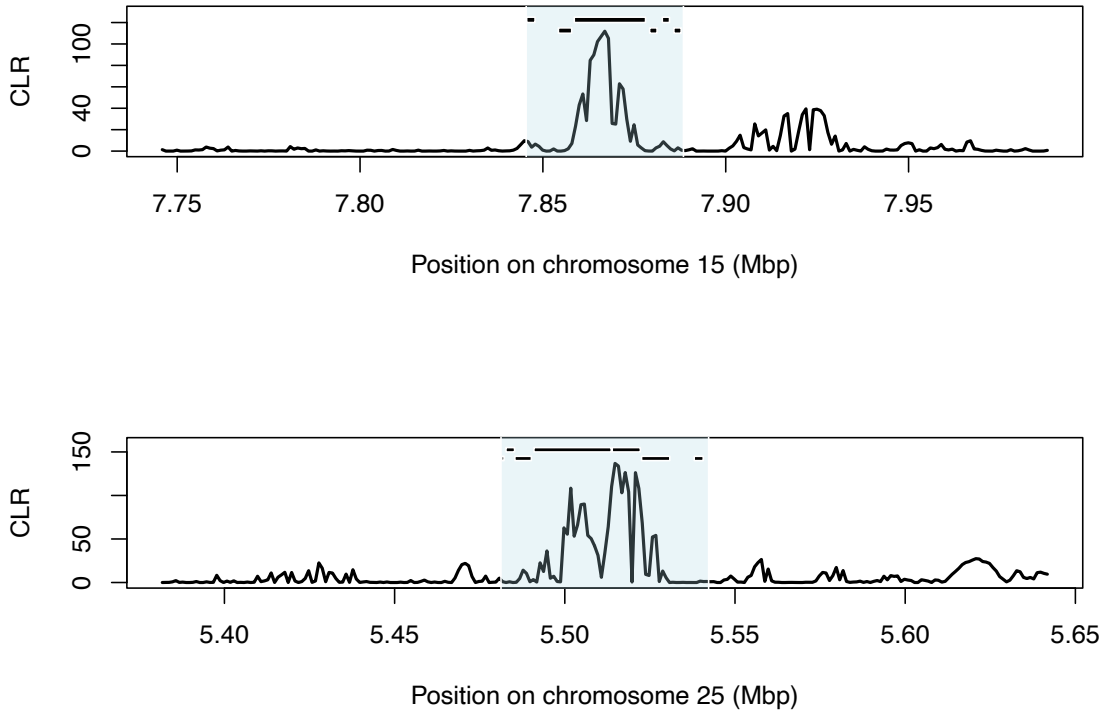


Figure 7: Sweep regions on chromosomes 15 and 25. Composite likelihood ratio (CLR) calculated in SweepFinder2. Black bars indicate the position of gene annotations on the forward strand (top) and reverse strand (bottom) listed in Supplementary Table S3. Blue shading indicates the sweep region.

Evidence of strong, recent selection at *CYP333B3*

The largest sweep occurred at the cytochrome P450 gene *CYP333B3* on chromosome 13. This gene is of interest because of its validated functional association with pesticide resistance in *Helicoverpa*, and because nonsynonymous substitutions in this gene are known to have increased in frequency in North American *H. zea.* Assuming direct selection at this locus, we estimate the selection coefficient $\hat{s} = 0.0489$ (Figure 8A). This estimate is based on estimates of the effective population size, the recombination rate, and qualities of the sweep (see Methods for discussion of our simplifying assumptions and sources of error).

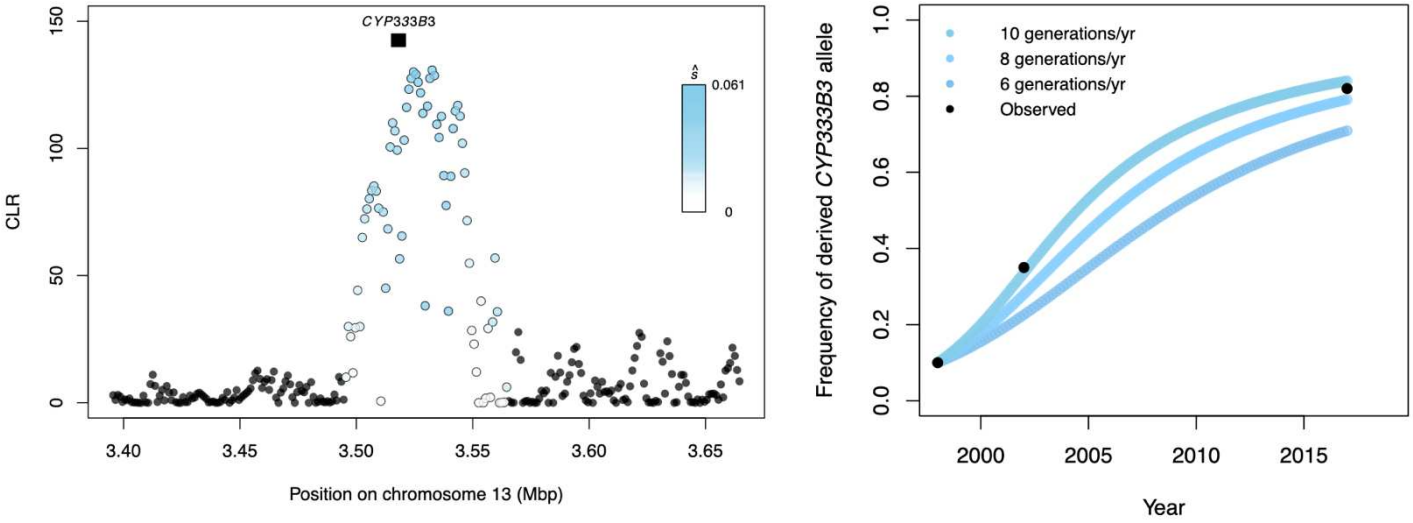


***Figure 8*: Selective sweep at *CYP333B3*.** **A**: Composite likelihood ratio calculated in SweepFinder2, with points coloured by their estimated selection coefficient. Black box indicates the position of *CYP333B3*. **B**: Predicted frequency of a dominant-acting derived *CYP333B3* allele for each generation given the estimated selection coefficient at that locus $\hat{s} = 0.0489$ assuming 6, 8 and 10 generations per year. Point colour corresponds to assumed generation time. Black points are independently estimated allele frequencies from Taylor *et al.* (2021).

In North American *H. zea*, Taylor *et al.* (2021) identified temporal genetic differentiation concentrated in a region that tightly overlapped with the sweep we observed on chromosome 13. In the same study, the authors were able to go back to *H. zea* samples from 1998, 2002 and 2017 in their freezer collection to genotype individuals at this locus, showing that the proportion of individuals with derived non-synonymous *CYP333B3* mutations had increased over time. This afforded us an opportunity to determine whether our estimate of the selection coefficient was consistent with the change in allele frequency that they observed.

Our estimate of $\hat{s}$ predicted the allele frequencies observed by Taylor *et al.* with less than 3% error when we assumed a dominant mode of inheritance and a biologically realistic generation time of 8-10 generations per year (Figure 8A, Supplementary Table S6). We also fit an estimate of the selection coefficient ($\hat{s}_{fit}$) to the allele frequencies empirically estimated by Taylor *et al.* and found that the two selection coefficient estimates differed by less than 0.05 for generation times between 5 and 10 generations/yr (Supplementary Table S6). The dominance coefficient that best explained the data was >0.997 regardless of generation time when jointly optimised with the selection coefficient, consistent with our assumption of complete dominance. Therefore, we were able to precisely approximate the strength of selection at this locus from a single-timepoint sample. Moreover, this approximation was accurate enough to infer that anthropogenic selection led to a fixation of a derived *CYP333B3* allele within ~20 years.

Patterns of genetic differentiation and diversity confirm that the selected allele arose within *H. zea* and was subject to selection within the last 20 years. When comparing *H. zea* samples from 2002 and 2019, genetic differentiation is greatest at the *CYP333B3* locus and otherwise low, consistent with a selective sweep occurring primarily in the intervening time (Supplementary Figure S6A). By contrast, differentiation is high across the full extent of chromosome 13 when comparing 2019 *H. zea* samples to *H. armigera* – especially at the *CYP333B3* locus (Supplementary Figure S6B). Among the 2019 samples, the average rate of coalescence was elevated at the locus. Specifically, there was a dearth of

genealogies coalescing >2N_e generations in the past at *CYP333B3*, (Supplementary Figure S6C) consistent with the effects of a selective sweep. These results contradict expectations under introgression (adaptive or otherwise), which would leave deep coalescence events at loci abutting the selected gene (Setter et al. 2020). On chromosome 13, the average height of coalescent genealogies estimated within individuals was 2N_e generations (standard error 1.5x10^{-4} N_e generations), contrary to what we would expect if some sampled haplotypes originated in a divergent population/species (see Methods). Together, these results indicate a strong selective sweep at *CYP333B3* over the past 20 years, and rule out introgression of *CYP333B3* from *H. armigera*.

Signatures of selection at candidate Bt-resistance genes

None of the candidate Bt resistance loci that we mapped occurred in sweep regions, however one candidate locus on chromosome 13 occurred within the upper 1st percentile of all CLR values (Figure 9). This locus (chr13:4012595-4041462) includes *PIK3C2A* and *kinesin-12*-like (positioned within *PIK3C2A*, on the opposing strand; Supplementary Table S4). Benowitz *et al.* (2022) identified a premature stop codon in *kinesin-12-like* (hereafter, *kinesin*-12) as the primary candidate target of selection underlying Cry1Ac resistance in the field-derived GA-R strain of *H. zea*.

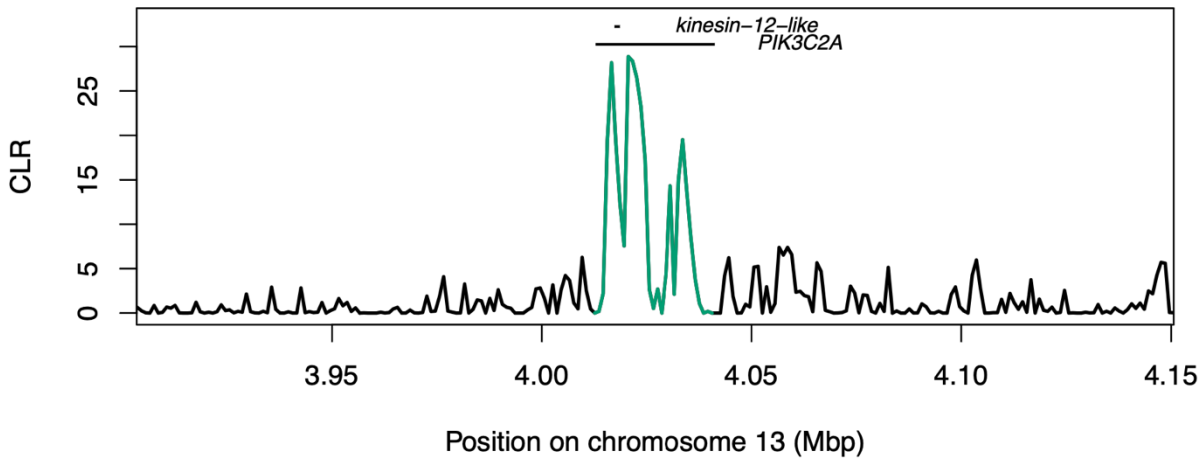


Figure 9: **Selective sweep likelihood ratio at the locus *PIK3C2A/kinesin-12*-like.** CLR values in the region of the loci *PIK3C2A* and *kinesin-12*-like, encoded on opposing strands. Both genes are candidate Cry1Ac resistance loci. Values overlapping with the annotation positions are highlighted in green.

Therefore, we investigated whether this same mutation occurred in our wild-caught samples. Requiring 95% confidence, we could call the presence or absence of the premature stop codon in 100 of our 237 samples. Of these, 99 carried the susceptible-strain allele and 1 carried a nonsynonymous C>A transversion mutating glutamine to lysine. We also identified 9 non-singleton SNPs in the coding region of *kinesin*-12 that were both nonsynonymous and resulted in amino acids with different biochemical properties to that of the susceptible strain (Supplementary Table S5). The most common of these was a C>A transversion mutating threonine to lysine, which occurred at allele frequency $q = 0.11$ (N=96 genotyped individuals). Therefore, the premature stop codon is unlikely to have caused the putative sign of selection we observed, though it is plausible that other nonsynonymous mutations have a similar phenotypic effect in disrupting the function of Kinesin-related protein 12. It is also possible that unobserved indels, larger structural variants, or cis-regulatory elements could cause the putative signal of selection observed at this locus. We note that the highest CLR values in this region align with segments immediately flanking *kinesin-12* (Figure 9).

Discussion

Gene flow within and between *Helicoverpa* species in North America

Our first aim was to test for the presence of invasive *H. armigera* ancestry in the north American *H. zea* population. Based on its rapid spread in South America, high propagule pressure, and availability of suitable contiguous habitat, Kritikos considered it was "a matter of time" until *H. armigera* – or introgressed *H. armigera* ancestry, at least – establishes in North America (Kriticos et al. 2015). Despite this, to our knowledge there has been no reported evidence of *H. armigera* in Central America. There has only been one report of *H. armigera* captured in the US prior to 2019, based on the COI haplotype of three individuals captured in Florida in 2015, with no evidence of establishment (Tembrock et al. 2019). We present the first conclusive report of admixed individuals in mainland North America. Only two individuals we sampled were admixed, and both were sampled at the same site in Texas despite an absence of population structure. The admixed samples show nowhere near the extent of *H. armigera* allele sharing observed in Brazil (Figure 1A). We saw no evidence of *H. armigera* introgression in previously reported samples collected in 2012 or 2017, though the sampling effort was far reduced in those years compared to our 2019 survey. Overall, the results point to a small degree of recent interspecific gene flow concentrated at a pesticide resistance locus.

Haplotypes break down over successive generations through recombination and mutation (Pool & Nielsen 2009). The size and therefore the detectability of introgressed ancestry blocks therefore also decays over time. Admixture between *H. armigera* and *H. zea* in South America was apparently punctuated by a pulse of hybridization 60-100 generations ago, with declining rates of hybridization since (Valencia-Montoya et al. 2020). We concluded that the vast majority of *H. zea* samples were non-admixed. Since we do not expect ABBA and BABA pattern frequencies to be exactly equal, values of f_d insignificantly greater than zero across several chromosomes (Figure 1B) are consistent with an absence of introgression. This does not rule out the possibility of rare, short *H. armigera* haplotypes segregating in the *H. zea* population. However, our method was sensitive enough to detect *H. armigera* introgression in a region spanning ~1% of the genome in <1% of individuals, suggesting that it was sufficient for our aim of identifying the presence of introgression at levels that are meaningful for the management of this pest.

Our second aim was to characterise population structure within non-admixed North American *H. zea*. While the use of genomic data has revealed otherwise cryptic population structure in some species, including *H. armigera* in its native range, we show that previous reports of complete panmixia in *H. zea* are not simply due to a lack of power (Zhang et al. 2022). Our analysis based on millions of SNPs showed that individuals collected at different locations in Texas were, on average, as genetically differentiated from one another as they were from individuals collected in North Carolina. While remarkable, this result is consistent with our understanding of *H. zea* movement ecology (Jones et al. 2018; Lawton et al. 2022). Given that summer incursions at latitudes higher than our sampling range are consistently repopulated by southern populations in the range we sampled, *H. zea* is likely panmictic across North America. Fitness-enhancing alleles can therefore spread readily across space. This highlights the need for monitoring and management at the species level across the agroecosystems potentially affected by *Helicoverpa*. Moreover, the evolution of pesticide or Bt resistance in one crop or region has immediate consequences throughout North America, where there may be substantially different agricultural and pest management practices (Reisig & Kurtz 2018). This result highlights the intrinsic ecological relationship between management practices in neighbouring fields, regions, or states, and should emphasise the need for multilateral management solutions.

Pesticide resistance through interspecific introgression

CYP337B3 could have been introduced via two routes: northward dispersal of admixed individuals from South America, or an independent introduction of *H. armigera* from its native range. An independent introduction is highly unlikely as it would require many generations of hybridization and recombination. Yet we saw no evidence of *H. armigera* ancestry in 2012 and 2017. Gene tree reconstruction could not be used to identify a clear source population of the *CYP337B3* polymorphism due to a lack of phylogeographic signal within *H. armigera* (Figure 5). However, the very fact that *H. armigera* ancestry is concentrated around the *CYP337B3* locus, which is overrepresented among admixed individuals in Brazil due to selection, (Valencia-Montoya et al. 2020) strongly suggests that the polymorphism was introduced through northward dispersal.

Our results show that the admixed samples carried alleles that, due to their semidominant phenotypic effect, enable resistance to fenvalerate, deltamethrin and possibly cypermethrin (Rasool et al. 2014; Ni et al. 2023; Joußen et al. 2012). Thus, an otherwise absent resistance phenotype was introduced to the North American *H. zea* population via interspecific introgression. The introgression of this gene is an indirect result of adaptation to pyrethroid exposure both in the native range of *H. armigera* (Rasool et al. 2014; Ni et al. 2023; Joußen et al. 2012) and subsequently in South American *H. zea* (Valencia-Montoya et al. 2020; Durigan et al. 2017). As anthropogenic activity increasingly brings pest species into secondary contact, introgression could become an important mechanism of rapid adaptation. This is particularly relevant to pesticide resistance, which can in some cases be underlain by a modular genetic architecture. For example, adaptive introgression of a *vkorc1* allele from the western Mediterranean mouse *Mus spretus* into the house mouse *M. musculus domesticus* confers the latter with resistance to anticoagulant rodenticides (Song et al. 2011). Introgression of an *ARH* allele from Atlantic to Gulf killifish species enabled tolerance to anthropogenic pollutants, which was shown to be highly fitness-enhancing (Oziolor et al. 2019). Similarly, interspecific introgression in the fungal genus that causes Dutch elm disease, *Ophiostoma*, was associated with virulence (Hessenauer et al. 2020). In each of these cases, genetic variation already shaped toward a fitness optimum in one species was introduced to another at relatively high frequency – an adaptive process that can be substantially faster than selection on *de novo* mutations (Marques et al. 2019).

Pesticide resistance through rapid adaptation in *H. zea*

Taylor *et al.* (2021) measured genetic differentiation in North American *H. zea* between 2002 and 2017, showing that a region overlapping with the sweep we identified on chromosome 13 was the most significant outlier. This result provides strong independent evidence that we have correctly identified a selective sweep. Two annotations lie within the sweep region: *carboxypeptidase Q*-like and the cytochrome P450 gene *CYP333B3*. *CPQ*-like is noteworthy because some carboxypeptidases bind Cry1Ac in *Helicoverpa armigera* (Da Silva et al. 2018) and are up-regulated upon Bt exposure in other lepidopteran pests (Yang et al. 2018; Van Munster et al. 2007), though to our knowledge have never been functionally associated with Bt resistance. Selection at *CYP333B3* is a more parsimonious explanation for the sweep. Cytochrome P450s have repeatedly been implicated in the evolution of resistance to a range of pesticides, especially in *Helicoverpa* (Nauen et al. 2022). The P450 gene *CYP333B3* plays a general role in xenobiotic metabolism in *Helicoverpa* and other noctuid crop pests (Amezian et al. 2021; Shi et al. 2022). Studies of noctuid pests have shown that the gene is induced by a broad array of pesticide classes: indoxacarb (a voltage-dependent sodium channel blocker), fluralaner (a GABA-gated chloride channel allosteric modulator), imidacloprid (a nAChR competitive modulator), aldrin (an organochloride), several host plant defences (xanthotoxin and gossypol), and the pyrethroid fenvalerate (Amezian et al. 2021; Shi et al. 2022; de la Paz Celorio-Mancera et al. 2011; Han Yangchun 2014). Two key observations link *CYP333B3* evolution to pesticide resistance in *Helicoverpa*. First, Shi *et al.* performed an in vitro metabolism assay in *H. armigera* and found that

583 *CYP333B3* showed the highest activity for metabolism of the organochloride aldrin (Shi et al. 2022).
584 Second, Han *et al.* showed that *CYP333B3* is constitutively overexpressed in the pyrethroid strain
585 FenR (Han Yangchun 2014). *CYP333B3* is therefore a key gene of interest with respect to pesticide
586 resistance.
587
588 We estimated that *CYP333B3* was subject to a selection coefficient $\hat{s} = 0.049$ and showed that this
589 coefficient could explain independently observed changes in the frequency of nonsynonymous
590 mutations from 0.07 in 1998 to 0.87 in 2017 (Taylor et al. 2021). *H. zea* resistance to pyrethroids was
591 first reported in 1990, and control failures became apparent in the South and Mid-West in the mid-
592 2000s when derived *CYP333B3* allele frequencies were approaching 0.4 (Jacobson et al. 2009). By the
593 late 2010s, pyrethroids were considered to have mixed efficacy compared to more effective diamide
594 pesticides (Olmstead et al. 2016). So although further functional studies are needed to confirm the
595 mechanistic link between the derived *CYP333B3* and pyrethroid resistance, two key observations lead
596 us to hypothesise that pyrethroid exposure has caused rapid adaptation at this locus over the past
597 two decades. First, there is clear evidence of a selective sweep at *CYP333B3* over the same narrow
598 window of time in which phenotypic change was observed in pyrethroid resistance in North American
599 *H. zea*. Second, there is a has been a detailed functional association between *CYP333B3* (its wild type
600 genotype in *H. armigera*, at least (Han Yangchun 2014)) and those same pyrethroids. Regardless of
601 the specific pesticide(s) that imposed positive selection on derived *CYP333B3* alleles, it is clear that
602 the adaptive response occurred within a ~20 year time frame – an example of rapid anthropogenic
603 adaptation. Although *CYP333B3* appears to be under convergent selection in *H. armigera* in Europe,
604 Africa and South America (Jin et al. 2023), we see no evidence of introgression into *H. zea* from *H.*
605 *armigera* at this locus. Rather, the increased rate of coalescence and genetic differentiation from *H.*
606 *armigera* and at this locus, combined with the recorded increase in frequency from a rare variant in
607 *H. zea* over a decade before the first observation of *H. armigera* in South America, strongly supports
608 selection from *de novo* mutation within *H. zea*. Thus, while *de novo* adaptation may be slower than
609 adaptive introgression, it can occur at timescales once thought to be impossible for an evolutionary
610 process. Therefore, for species with large and highly connected populations exposed to strong
611 selection pressures, management strategies must not only consider ecological impacts but also
612 evolutionary change over the short- and medium-term.
613
614 ## Other signatures of selection in North American *H. zea*
615
616 Bt resistance in *H. zea* is a major concern. Extensive research has shown that resistance to some
617 toxins (*e.g.* Cry1) is already common, and that selection for resistance to Cry2Ab2 was ongoing at the
618 time we collected samples (Bilbo et al. 2019; Ali et al. 2006; Yu et al. 2021; Reisig et al. 2018; Dively et
619 al. 2016). Although phenotypic data clearly show that resistance is evolving, identifying its genetic
620 basis in both the field and the lab has been challenging. This is because the same phenotype can
621 result from many polygenic, semi-overlapping genetic architectures. Multiple recent studies have
622 identified Bt resistance QTL that do not overlap with any established Bt resistance loci (Taylor et al.
623 2021; Benowitz et al. 2022). Detecting the signature of selection on such complex traits is difficult
624 without extensive replication or time-series data. Taylor *et al*. (2021) identified several QTL associated
625 with Bt resistance but found no evidence of allele frequency shifts at these loci over time. It is
626 perhaps not surprising, then, to find that none of the 19 Bt resistance QTL we investigated stood out
627 as showing clear selective sweeps. One notable exception is the locus *PIK3C2A*/*kinesin-12* (Figure 9),
628 in the upper first percentile of CLR values genome-wide. Benowitz *et al.* (2022) identified a premature
629 stop codon in *kinesin-12* as the primary Bt resistance candidate in their QTL. We did not observe this
630 mutation in our field-collected samples, though we did see non-singleton, nonsynonymous mutations
631 in the coding region of *kinesin*-12 that altered the physio-chemical properties of amino acids and may
632 have had the same phenotypic effect on disabling protein function. Cis-regulatory mutations could
633 also disable the function of Kinesin-related protein 12; consistent with this hypothesis, the highest

CLR values abut *kinesin*-12. So while the specific mutation observed by Benowitz *et al.* likely does not explain the putative signature of selection we observe, it is plausible that there has been selection at this locus in the field. Two important caveats should be noted. First, we could confidently call genotypes in ~40% of our samples, so we cannot rule out selection on the premature stop codon itself in the field. Second, this locus is only noteworthy because of our specific candidate gene search; it would not have been included amongst the most likely targets of interest in our selective sweep analysis alone. Even though it is within the upper first percentile of CLR outliers, CLR values are distributed such that hundreds of other genes show more obvious signatures of selection (Figure S5). Therefore, functional work is needed to test Benowitz's hypothesis that a premature stop codon in *kinesin-12*-like contributes to Bt resistance in the field.

To define sweeps in our hypothesis-free selection scan, we used a stringent yet arbitrary (upper 0.01st percentile of CLR values) in order to avoid false positives, though this means that real selective sweeps may be missed. For example, chromosomes 10 and 26 showed CLR values only 4 units below the threshold (Figure S5). For two of the three putative sweeps we identified, we can only speculate on possible causes of selection. The sweep on chromosome 15 includes a *takeout*-like gene. *Takeout*-like genes are over-expressed in pyrethroid resistant mosquitos and aphids; in the latter RNAi experiments showed that they directly contribute to resistance (Peng et al. 2021; Toé et al. 2015). *Takeout*-like genes were also significantly over-expressed in *Spodoptera litura* upon exposure to the isoxazoline insecticide fluralaner, and in honeybees exposed to the herbicide atrazine (Jia et al. 2020; Wang et al. 2023). It is therefore possible that this sweep is due to selection on a cis-regulator of *takeout*-like, though showing this requires further investigation. It is also noteworthy that the chromosome 15 sweep overlaps with two genes encoding E3 ubiquitin-protein ligases involved in protein degradation and cell cycle regulation (*Hyd*-like and *RNF168*-like) (Flack et al. 2017). Both *Hyd* and *RNF68* have been implicated in host-virus protein-protein interactions, consistent with selection imposed by a viral outbreak (Lilley et al. 2010). Some of the clearest signatures of selection result from outbreaks of infectious disease (Obbard et al. 2011). This may be particularly likely in North American *H. zea*, where nucleopolyhedroviruses are often used as a means of biocontrol alongside Bt toxins and synthetic pesticides (Niedermann et al. 2017).

The cause of the putative sweep on chromosome 25 is less clear, though we note that the dynein gene in this region was found to be the most up-regulated gene in an RNA-seq comparison of neonicotinoid-susceptible vs. resistant honeybees (Bahia 2021). While there is insufficient *a priori* information to go beyond a simple description of the sweeps on chromosomes 15 and 25, the plausible cause of selection on chromosome 13 is more obvious.

Pleiotropy, epistasis, and the evolutionary fate of *CYP337B3* and *CYP333B3* alleles

We identified two cytochrome P450 genes recently introduced to the *H. zea* metapopulation via completely distinct evolutionary processes. We have argued that both are associated with pyrethroid resistance. This begs the question: is there an epistatic interaction between *CYP333B3* and *CYP337B3*, and if so, what does this mean for the future spread of *H. armigera* ancestry? If the derived allele of *CYP333B3* is common and *CYP337B3* is rare, is the marginal fitness of the latter completely diluted in this population? Several factors need to be considered. First, resistance to a specific pesticide is a quantitative trait for which distinct cytochrome P450 genes often have additive effects (Brun-Barale et al. 2010; Tchouakui et al. 2021). Second, pyrethroids represent a vast and diverse set of pesticides. No single mutation confers complete resistance to all pyrethroids, and different mutations associated with pyrethroid resistance almost always contribute to resistance in a nonoverlapping set of pesticides. This can be the case for completely different classes of pesticide. In other words, both resistance alleles have pleiotropic phenotypic effects. Third, the derived *CYP333B3* allele was already common in North America when *CYP337B3* was introduced to South American *H. zea* populations in

2013 (Figure 8B). *H. zea* populations are well connected between North and South America, as evidenced by the spillover of *H. armigera* ancestry that we report here. Pyrethroids have been commonly used in both Brazil and the US for pest control in cotton and maize (Walsh et al. 2022). Given that *CYP337B3* was highly fitness-enhancing in South America, and given that it would likely have often occurred in a genomic background with this same *CYP333B3* allele, we have little reason to think it would not spread in North America as well. That said, a fourth factor to consider is demography. Due in part to the phase of the El Niño southern oscillation, a population expansion of *H. armigera* in Brazil shortly after its detection appears to have resulted in demographic swamping, producing a pulse of hybridization (Specht et al. 2021; Valencia-Montoya et al. 2020). By contrast, in 2019 we see a 'trickle' of *H. armigera* ancestry into North American *H. zea*. Even alleles under strong selection can be lost through drift at sufficiently low frequency. Therefore, the ongoing influx of *H. armigera* ancestry and the North American selective regime with respect to *CYP337B3* will together determine spread. While a degree of predictive power could be gained through fitness assays to determine the epistatic interaction between *CYP337B3* and *CYP333B3* for tolerance to fenvalerate and other commonly used pesticides, the outcome of this incursion remains to be seen through future biosurveillance.

Prospects for genomic surveillance in this system and others

Early detection of invasive pests can pay dividends even when surveillance is costly, as invasive species are easier to manage shortly after they establish (Mehta et al. 2007). Genomic detection of *H. armigera* is currently far more scalable and sensitive as a diagnostic test to distinguish *H. zea* from *H. armigera* relative to morphological or single-locus genetic assays. Nonetheless, a trade-off exists between the cost of surveillance and the product of the probability and cost of invasion. Given that *H. armigera* ancestry has spread into the United States, and the value of the crops exposed to *H. armigera* (~US$78 billion per annum (Kriticos et al. 2015)), there is a strong case for extensive ongoing genomic monitoring of *H. zea* across North America, especially in the dispersal corridor leading through Central America and Mexico where there has been comparably little genomic monitoring to our knowledge. Monitoring programs must adopt a genic view of biological invasion in order to quantify the incursion of *H. armigera* (North et al. 2021). This means that, if the aim is to map the geographic distribution of possible *H. armigera* spread, *CYP337B3* assays are likely to be more informative than ancestry-based assays that use one or few markers. Alternatively, if the aim is to determine the presence of *H. armigera* introgression in an area of concern, only genomic data offer the resolution required to detect the short haplotypes segregating in the population. Future work to characterise the sets of loci most sensitive to admixture (*e.g.* those in high-recombining regions, least likely to be removed by selection due to linkage with Bateson-Dobzhansky-Muller incompatibilities) will allow for the development of SNP arrays that can robustly assign ancestry proportions to samples at lower cost.

When genomic data was first adopted in the field of molecular population genetics, there was excitement about the power of population genomics to infer the evolutionary history and genetic basis of adaptive traits in wild populations, especially where laboratory crosses were impossible (Li et al. 2008). While genomic surveillance has become a critical technology for the management of pests and invasive species (Hamelin & Roe 2020; Chown et al. 2015), in practice the detection of adaptation can be limited by the complex genetic architectures and demographic nonequilibrium. Incorporating QTL metadata and sampling over a time-series can address these issues – reinforcing the need for sustained genomic monitoring programmes – though this is not always possible (Taylor et al. 2021; Clark et al. 2023; Pélissié et al. 2018). Although we only sample individuals at a single timepoint, we show that relevant population genetic statistics can be estimated with sufficient accuracy to not only identify a putative target of selection for pesticide resistance, but to precisely infer recent allele frequency change at the locus. The analyses we apply here demonstrate the richness of information

that can be extracted from nucleotide diversity in wild pest populations to inform management action and study anthropogenic adaptation.

In summary, we have shown that pesticide resistance in North American *H. zea* arose recently, and rapidly, via two independent processes: interspecific introgression and intraspecific adaptation. These findings underscore the importance of rapid adaptation for pest management – evolution at timescales previously only considered to be relevant for ecological process.

Methods

Sampling, DNA extraction, library preparation and sequencing

H. zea individuals were collected in agricultural fields at 10 locations across the southern U.S. in 2019 (see Supplementary Table S1 and Figure 1) using corn earworm pheromone lure (GreatLakes IPM, Vestaburg, MI). Adult moths were immediately placed in 95% ethanol and stored at -20°C prior to DNA extraction. A few populations of *H. zea* were collected as larvae (see Supplementary Table S1) and brought back to the laboratory of Dr. David Kerns at the Texas A&M University Dept. of Entomology in College Station, TX, USA where they were raised to adults and then processed for DNA extractions.

The abdomen of each adult moth was separated from the body and used for DNA extraction with the Qiagen DNeasy Blood & Tissue kit (Qiagen, Germantown, MD following the manufacture's protocol. The final DNA elution step was performed using Qiagen buffer EB (Qiagen, Germantown, MD) instead of the AE buffer. DNA concentration was measured using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA). Genomic DNA samples were delivered to the Texas A&M AgriLife Genomics and Bioinformatics Service in College Station, TX for library preparation and sequencing.

DNA batch purity and integrity were assessed using a DS-11 spectrophotometer (Denovix, Wilmington, DE) and capillary gel electrophoresis (Fragment Analyzer, formerly Advanced Analytical, now Agilent, Santa Clara, CA), respectively, from 11 samples per 96-well plate. Genomic DNA was further purified with SPRI beads (Omega Bio-tek, Norcross, GA) and the concentration of each sample was determined using Lunatic plates (Unchained Labs, Pleasanton, CA) read on a DropletQuant spectrophotometer (PerkinElmer, Waltham, MA).

Libraries were prepared from 33ng of genomic DNA using a custom miniaturized version of the NEXTFLEX Rapid XP kit protocol (PerkinElmer) that was automated on a Sciclone NGSx liquid handler (PerkinElmer). Briefly, all reactions involving enzymes were carried out in one third of manufacturer proscribed volumes for each reagent. In the modified reaction volumes, genomic DNA was enzymatically fragmented for 3 minutes, then ligated to unique dual-indexed barcodes. The raw library reaction was brought up to protocol volumes with H_2O for a SPRI cleanup followed by SPRI size selection between 520 and 720 bp, and then eluted in reduced volume to be amplified for 10 PCR cycles. Finally, the amplification reaction was brought back up to recommended volume with H_2O to perform an additional single sided SPRI size selection that retains DNA fragments larger than 450bp.

A subset of libraries was checked for size and integrity on the Fragment Analyzer and all libraries were quantified using a fluorescent plate reader (SpectraMax M2, Molecular Devices, San Jose, CA) with PicoGreen reagent as per the manufactures suggested protocol (Thermo Fisher Scientific, Waltham, MA). Libraries were diluted with EB (Omega Bio-tek) to a final concentration of 2.25 ng/µl using an automated liquid handler (Janus, PerkinElmer) and an equal volume of each was pooled. Pooled library quality was assessed on the Fragment Analyzer and molarity determined using a qPCR-based Library Quantification assay (Roche, Pleasanton, CA).

The pool was sequenced in a single lane of an Illumina NovaSeq S4 XP flowcell (San Diego, CA) using the 2 X 150 bp recipe. The raw data was demultiplexed with bcl2fastq 2.20, which yielded 2.17 billion demultiplexed reads ranging from 5.01-10.41M reads per sample and an average of 7.80M reads per sample (~6.5x coverage).

Filtering, alignment, and variant calling

In order to conduct intra- and inter-specific comparisons, we used raw sequencing data from multiple *Helicoverpa* spp. generated by Anderson *et al.* (2018), Taylor *et al.* (2021) and Jin *et al.* (2023). These publicly available data were analysed in the same manner as our newly generated sequence data. Two slightly different pipelines were used for different analyses depending on the samples required. Bioinformatics pipelines were implemented using Snakemake v7.2 (Mölder et al. 2021).

For the results presented in Figures 1-6 (hereafter, Call Set 1): Fastq files were trimmed using fastp v0.23 (Chen et al. 2018), then mapped using bwa v 0.7.12 (Li & Durbin 2009). The resulting .bam files were sorted using SAMtools (Li et al. 2009). Picard v2.9.2 (Broad Institute 2023) was used to remove duplicates and SAMtools was used to index the filtered .bam files. GATK HaplotypeCaller v4.3 (van der Auwera et al. 2013) was used to call haplotypes per individual. GATK CombineVCFs was used to merge call sets across all individuals and GenotypeGVCFs was used to jointly call genotypes across samples sequenced here, those produced by Anderson *et al.* and samples collected in 2002 by Taylor *et al.* (2021). Of these, only single nucleotide polymorphisms were retained. VCFtools (Danecek et al. 2011) was used to filter out sites with a phred-scaled quality score below 20, sites with a mean depth of coverage below 2X or above 200X, genotypes with a mean depth below 2X or above 200X, with a missing data threshold of 50%.

These samples were mapped against the highly contiguous *H. armigera* reference genome to maximise sensitivity when detecting introgression from *H. armigera*, and to allow direct comparison with results generated in similar studies (Taylor et al. 2021; Valencia-Montoya et al. 2020). The *H. armigera* genome is highly contiguous with the *H.* zea chromosome except for the sex chromosome (Benowitz et al. 2022). Therefore, some *H. zea* sex chromosome or pseudo-autosomal reads miss-mapped to autosomal scaffolds in the reference assembly. To avoid potential biases introduced through miss-mapping, Plink v1.9 (Chang et al. 2015) was used to calculate the statistical association between each autosomal SNP and heterozygosity on the Z chromosome, in which Z chromosome heterozygosity was treated as a quantitative trait. SNPs with *p* values in the top first percentile of association were excluded from all downstream analyses. This resulted in a call set of 45,700,390 SNPs among 304 individuals of four species (Supplementary Table S2). For analyses requiring additional samples (*e.g. H. zea* from 2017 and 2012, and those shown in Figure 5) mapping and variant calling was completed a subset of relevant samples was jointly called using the same pipeline.

For the selection analysis (results presented in Figures 7-9; hereafter Call Set 2) a more stringent filtering pipeline was used. Paired end raw reads were trimmed using Trimmomatic v0.38 (parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50) (Bolger et al. 2014). Trimmed and paired reads were aligned to the *H. armigera* reference assembly generated by Pearce *et al.* (Pearce et al. 2017). Alignment was performed using bwa (Li & Durbin 2009) and mate pairs were fixed using SAMtools v1.9 (Li et al. 2009). Resulting bam files were sorted using SAMtools. Picard v2.18 was used to add read groups, clean bam files, and mark duplicates. Subsequently, SAMtools was used to index the bam files, and GATK HaplotypeCaller v4.15 (van der Auwera et al. 2013) was used to call haplotypes individually for each sample. Haplotypes were imported into a database using GenomicsDBImport from GATK. Lastly, the module of GenotypeGVCFs from GATK was used to call SNPs from all samples jointly. Next, the called genotype in a vcf file was subjected to a sequential

filtering process using VCFtools v0.1.16 (Danecek et al. 2011). First, we removed samples that had over 50% of loci missing; these samples usually resulted from poor sequencing coverage. Second, we applied a quality filter by removing loci that had a genotype quality score lower than 20 (minGQ 20). Third, we applied a depth filter to remove loci that had coverage less than 2 or greater than 200. We only kept biallelic SNPs and removed all indels. Next, we removed loci missing from more than 50% of the samples missing, as well as all singletons in the vcf file. In the next step, we filtered out loci that violated Hardy-Weinberg equilibrium (HWE). HWE filtering was only conducted for *H. armigera* and *H. zea* samples because of the much smaller sample size of other species. HWE tests were applied only to loci with no missing genotypes, and *p*-value cutoff was set to 0.01. Loci that violated HWE in two of the three species were removed. We imputed the missing genotypes using Beagle (v5.1, Browning et al.) with default settings. Then we only kept biallelic SNPs in the imputed genotypes, and we also applied linkage disequilibrium pruning using Plink (--indep-pairwise 50 10 0.5). SNPs associated with Z chromosome heterozygosity were filtered using the method described above, resulting in a set of 5,706,238 SNPs across 237 *H. zea* individuals.

For most window-based analyses described below, a window size of 20kbp was chosen because linkage disequilibrium decayed to a genome-wide background level of r^2 = 0.002 over approximately 20kbp for the population of North American *H. zea* that we sampled.

Detecting introgression

To test for allele sharing with South American *H. armigera,* we calculated $\widehat{f_d}$ in 20kbp windows across all chromosomes in which at least 200 SNPs were called, using Python scripts described in Martin in *et al.* (Martin et al. 2015). P1 was designated as a set of 13 *H. zea* individuals sampled in 2002 in Louisiana by Taylor *et al.* (2021). These samples were collected over a decade before *H. armigera* was detected in the Americas and therefore represent the best possible whole-genome reference set of non-admixed *H. zea*. Note that we do not use *H. zea* individuals sampled in New York in 2005 from Anderson *et al.* (2018), as the metadata associated with these samples was contradictory and could not be verified. P3 was 25 non-admixed *H. armigera* sampled in Brazil, and the outgroup was 7 *H. punctigera* individuals sampled in Australia. For our samples, $\widehat{f_d}$ was calculated 10 times where the test set, P2, was *H. zea* sampled in 2019 at each of the 10 samples sites. And additional three tests were carried out *H. zea* collected in 2012 in Louisiana, and collected in 2017 in both Maryland and Louisiana by Taylor *et al.* (2021). As a positive control, $\widehat{f_d}$was also calculated for 9 individuals sampled in Brazil shown to be admixed offspring of *H. armigera* and *H. zea* by Anderson *et al.* (Anderson et al. 2018). Results from all 14 tests are shown in Figure 1.

Calculation of summary statistics and visualisation with PCA

To compare the two admixed individuals with *H. armigera* samples we calculated nucleotide diversity (π), genetic differentiation (F_{ST}), and genetic divergence (d_{xy}) for the results presented in Figure 3 and Supplementary Figure S2 using python scripts described by Martin *et al.* (Martin et al. 2015). This was done in 20kbp and 100kbp windows (with at least 200 and 100 informative sites, respectively) with consistent results. Principal components analysis (Figure 4) was carried out using Plink v1.9 (Chang et al. 2015) in segments A and B separately.

Reconstructing gene trees

To reconstruct a tree at the *CYP337B3* locus, we used a combination of *H. zea* samples collected here (two admixed and two representative non-admixed samples), 4 *H. zea* collected in 2002 by Taylor *et al.* (2021), 2 *H. punctigera* and 9 *H. armigera* collected from both the invasive range in Brazil by Anderson *et al.* (2018) (which are admixed with *H. zea*), and 10 *H. armigera* from the major clades of

the native range by Jin *et al.* (2023) (see Supplementary Table S2). These samples were mapped, jointly genotyped and filtered without phasing in the manner as described above. We used VCFtools v0.1.15 (Danecek et al. 2011) to subset the *CYP337B3* locus previously used to identify signatures of adaptive introgression (HaChr15:11436565-11440168; Valencia-Montoya et al. 2020) from a multigenome VCF and converted to PHYLP format. We removed invariant and uninformative sites and retained sequences with no more than 50% missing data. We reconstructed a maximum likelihood phylogram the GTR+GAMMA model of rate heterogeneity and Lewis ascertainment bias correction method implemented in RAxML v 8.2.12 (Stamatakis 2014) with 100 bootstrap iterations. We rooted the tree using the two outgroup *H. punctigera* samples in ggtree (Xu et al. 2022), ensuring bootstrap values were assigned to the correct node with the edgelables function (Czech et al. 2017). The resulting cladogram is shown in Figure 5. The bootstrap values, tip label sample names and scaled branch lengths are shown as a phylogram in Supplementary Figure S4, which was rooted in FigTree v1.4.4.

Measuring Isolation by Distance

To investigate connectivity across the North American *H. zea* metapopulation, we calculated the mean of F_{ST} across 20kbp windows from all chromosomes for each of 45 possible pairwise combinations of the 10 sampling locations. Geographic distance was calculated for the same pairwise combinations using the R package geosphere (Karney 2013). We regressed $\frac{\overline{F_{ST}}}{(1-\overline{F_{ST}})}$ on log-transformed geographic distance to test for a positive correlation. A positive correlation would suggest a pattern of isolation-by-distance and provide a rough estimate of the product of population density and the variance in dispersal distance, whereas the absence of a correlation would indicate panmixia across sampling locations (Rousset 1997).

Identifying selective sweeps

Briefly, we identified selective sweeps as localised and extreme deviations from the genome-wide site-frequency spectrum (SFS) consistent with linked selection using the composite likelihood ratio test implemented in *SweepFinder2* (DeGiorgio et al. 2016; Nielsen et al. 2005). To do so, we used sequence data from other *Helicoverpa* species to distinguish ancestral from derived alleles. By using the empirical SFS as a null model, rather than an equilibrium null SFS, we account for the confounding influence of demographic nonequilibrium (Nielsen et al. 2005). We use sequence data from other *Helicoverpa* species to estimate chromosome-wide mutation rates in order to estimate the population recombination parameter, then use chromosome-wide recombination rate estimates and estimates of the effective population size to estimate the strength of selection that acted at a locus of interest. Finally, we show that the estimated selection coefficient can explain independently observed shifts in allele frequency.

Specifically, we used VCFtools v0.1.15 (Danecek et al. 2011) to calculate allele frequencies within the North American *H. zea* population sampled in 2019. Including 25 non-admixed *H. armigera* individuals and 7 *H. punctigera* individuals as outgroups, we used a custom python script to designate ancestral and derived states to those alleles and calculate all possible transition and transversion rates for each chromosome. With these ancestral and derived states, *SweepFinder2* (DeGiorgio et al. 2016) was used to compute the unfolded genome-wide site frequency spectrum in the North American *H. zea* population, and to use this spectrum as a null model in a scan for selective sweeps. Briefly, *SweepFinder2* runs a composite likelihood ratio (CLR) test to identify chromosomal regions that show a decline in genetic diversity characteristic of genetic hitchhiking (Kim & Stephan 2000). The CLR was computed in 1kbp intervals. Note that the LD-based local recombination rate estimates generated below were not used to detect the location of selective sweeps, as this would have

confounded the localised effect of selection on both the site-frequency spectrum and on linkage disequilibrium. We identified sites of interest (those clearly showing signs of selection) as the top 0.01st percentile of CLR values (CLR>104.17). Sites above this threshold within 20kbp of one another were grouped into clusters using BEDtools v2.20.1 (Quinlan & Hall 2010), whereby the cluster start and end positions extended 20kbp up/downstream of the terminal outlier sites. We refer to these regions as putative selective sweeps.

Mapping known candidate Bt resistance loci

We identified three classes of previously identified candidate Bt resistance loci: genes within the QTL described by Benowitz *et al.* (2022); genes repeatedly implicated in Bt resistance which were also examined by Taylor *et al.* (2021); and the novel resistance QTL identified by Taylor *et al.* (2021).

First, Benowitz *et al.* (2022) mapped a 250kbp Bt resistance QTL to chromosome 13 containing 10 genes of interest, including the putative causative locus *kinesin-12*-like, using a different reference assembly. To test for signatures of selection in this region, we used BLAST v2.4 to map each of the 10 loci in Table 2 of Benotitz *et al.* to our reference assembly, retaining the largest region of overlap with an *E*-value of 0 for each locus. As expected given the contiguity between our assemblies, all loci mapped within 250kbp of one another on chromosome 13. Within each BLAST hit, we used BEDtools v2.20.1 BEDops 2.4.41 to identify annotations in our assembly with matching annotation names (Quinlan & Hall 2010; Neph et al. 2012). Two annotations (*kinesin-12*-like, *JHE*-like) were absent in our annotation; we excluded the latter because we could not confirm its location. Because *kinesin-12*-like mapped well to our assembly (>90% sequence identity, *E*-value=0), and since *kinesin-12*-like sits within the larger gene *PIK3C2A* (where it is encoded on the opposite-sense strand) we could map the 1.7kbp gene with a range of error <50bp.

Second, we applied the same mapping approach to identify other known Bt resistance loci listed in Taylor *et al.* (2021) Table S9. We were able to confidently assign corresponding annotations for all genes except alp, cad2, *calp4* and *abcA2*.

Third, Taylor *et al.* (2021) identified many QTL associated with resistance to crops expressing both Cry1Ab and Cry1A.105 plus Cry2Ab2 toxins. Both traits were highly polygenic. We identified the most significant outlier scaffold for trait for both traits and mapped these to our assembly. This was done for separately for *H. zea* scaffolds and *H. armigera* super-scaffolds. To be consistent with Taylor *et al.* as much as possible, we defined the most significant outlier scaffold as that with the lowest mean *p* value for the linear mixed model likelihood ratio test with at least one SNP assigned a2 BSLMM posterior inclusion probability >0.01. We retained only super-scaffolds with at least 10 SNPs and scaffolds with at least 3 SNPs. We identified the scaffold NW_018395566 and the nested scaffold KZ118765 (nested within NW_018395566) as the scaffolds most associated with Cry1Ab resistance NW_018395399 and the nested KZ118015 were most associated with Cry1A.105+Cry2Ab2 resistance. All four were within the resistance-associated linkage group 9 in the map produced by Taylor *et al.* and neither were associated with growth rate on either of the control treatments. We mapped these scaffolds to our reference assembly, retaining the largest hit with an *E*-value of 0 as above.

For all candidate Bt-resistance loci that we mapped, we performed a permutation test to determine the probability that the locus overlapped with outlier sweep CLR values by chance. To do so, we randomised the position of a locus of the same size in order to generate a null distribution of overlapping mean and maximum CLR values (Supplementary Table S4).

Calling *Kinesin-12* genotypes

To directly compare *kinesin*-12 genotypes to those reported by Benowitz *et al.* (2022), we mapped reads to the full *kinesin*-12 sequence (including untranslated regions and introns) of the LAB-S (laboratory susceptible) *H. zea* strain. This was done to rule out the potential effect of alignment errors on genotype calls and enable an unbiased comparison of genotypes with our wild-caught individuals. This is because the *H. armigera kinesin*-12 sequence differs structurally from that of *H. zea* in non-coding regions, though its position in the genome is concordant. For all 237 2019 samples, the same pipeline applied to call set 1 was used to process fastq files, map reads and call variants with HaplotypeCaller. Three different parameters were used: the LAB-S *kinesin*-12 sequence was used as the reference sequence for mapping, no more than two alternate alleles were permitted per site, and monomorphic sites were retained. When calling variants, a cut-off of 95% confidence was used (Reference genotype quality or QUAL > 13.0103 for monomorphic and polymorphic sites, respectively). Putative indels were not considered. The resulting sequences were translated using EMBOSS Transeq and aligned (along with the coding sequence of the *Bt*-resistant strain) using Mview (Madeira et al. 2022)This alignment confirmed that the *Bt*-resistant line carried at C>T mutation resulting in a premature stop codon, and allowed us to directly compare our samples.

We characterised nonsynonymous mutations of interest that could be confidently called in our samples. We reasoned that amino acid changes with novel biochemical properties are more likely to impact enzymatic function, that singleton mutations are more likely to be erroneous genotype calls, and that singleton mutations are less likely to occur at sufficient frequency to cause the putative signature of selection in the region. Therefore, we only report nonsynonymous mutations (relative to the coding region of the reference LAB-S complete sequence) that (1) produce amino acids with different biochemical properties, and (2) could be confidently called in more than one individual. Only single nucleotide polymorphisms (homozygous and heterozygous) were reported (Supplementary Table S5).

Estimating the selection coefficient, s

For each site in which the CLR was calculated above, SweepFinder2 was also used to calculate Durrett and Schweinsberg's (2004) approximation of α, where

$$s = \frac{r \,.\ln(2N_e)}{\alpha}$$

This relationship allowed us to generate an estimate of the selection coefficient $\hat{s}$ in the sweep region on chromosome 13. To do so, we estimated the effective population size of our North American *H. zea* samples as $N_e = \frac{\hat{\theta}_\pi}{4\mu} = 10.3\text{x}10^5$, using the *Drosophila melanogaster* mutation rate μ = 8.4x10^{-8} (Haag-Liautard et al. 2007) and $\hat{\theta}_\pi = 0.0346$, calculated as the mean nucleotide diversity in 10kbp windows using pixy (Korunes & Samuk 2021). This effective population size is approximately half that of native *H. armigera* populations and consistent with previous estimates (Anderson et al. 2018). We used the mean estimate of per-base pair recombination rate for chromosome 13 (see below): $r =$ 5.447x10^{-8}. The mean selection coefficient within the bounds of the *CYP333B3* locus was $\hat{s} = 0.0489$.

Estimating chromosome-wide recombination rates

Estimation of the selection pressure, above, required mean per-chromosome estimates of the recombination rate. We estimated recombination rates across all chromosomes using patterns of linkage disequilibrium. First, we statistically phased data across all the *H. zea* individuals we sampled for all chromosomes using Beagle v5.0 (Browning & Browning 2007, 2016). We used 50 individuals with the highest depth of coverage to estimate the population recombination parameter ρ across all autosomes with LDhelmet v1.9 (Chan et al. 2012). We generated likelihood lookup tables using the same theta estimator described above, with the grid of ρ values between 0 and 10 specified with the command -r 0.0 0.5 3.0 1.0 10.0. Following the recommendation of Chan *et al.* (26), 11 Padé coefficients were computed and, using a default block penalty of 50 and a window size of 50 SNPs, the Markov Chain Monte Carlo procedure was implemented for 10^6 iterations with a burn-in of 10^5 iterations. To compare observations with other data and confirm that chromosome-wide estimates of recombination rate were biologically realistic, we converted mean estimates of $\rho = 4N_e c$ (for a recombination rate c per bp), to units of centimorgans per bp: $cM = 50 \ln \frac{1}{1-2c}$. As expected, recombination rate was negatively correlated with chromosome length (Supplementary Figure S7), the gradient of this relationship was similar to that observed in *Heliconius* butterflies, and the range of recombination rates observed overlapped with those measured in both *Heliconius* and *Drosophila* (Chan et al. 2012; Martin et al. 2019).

Modelling selection on *CYP333B3*

Taylor *et al.* (2021) genotyped *CYP333B3* from *H. zea* specimens sampled in the US at in 1998, 2002 and 2017, reporting an increase in the proportion of individuals with a derived genotype over time. We were therefore interested to know whether our estimate of $\hat{s} = 0.0489$ could explain these independently estimated shifts in the derived allele frequency, q. Starting from an observed initial allele frequency of $q_{1998} = 0.1$, we modelled the change in allele frequency in each subsequent generation as

$$\Delta p = \frac{pq.[p(w_{AA} - w_{Aa}) + q(w_{Aa} - w_{aa})]}{\bar{w}}$$

Here the reference genotypic fitness is homozygous ancestral, *i.e.*

$$w_{AA} = 1$$
$$w_{Aa} = 1 + h\hat{s}$$
$$w_{aa} = 1 + \hat{s}$$

So the mean fitness is

$$\bar{w} = p^2 w_{AA} + 2pq w_{Aa} + q^2 w_{aa}$$

We predicted allele frequencies through time assuming a dominant-acting *CYP333B3* mutation (h = 1; Figure 8B) and assuming codominance (h = 0.5; Supplementary Figure S8) for a range of generation times between 2 and 10 generations per year. The biologically realistic range of generation times in this case is unlikely to be below 5 generations/yr (Hardwick 1965; Morey 2010; Parajulee et al. 2004).

Additionally, we compared our estimate $\hat{s}$ to the selection coefficient that best explained the allele frequencies observed by Taylor *et al.*, $\hat{s}_{fit}$ .To estimate $\hat{s}_{fit}$, we ran 10^6 iterations of the model with a random selection coefficient, retaining the coefficient that maximised the fit to the data (the fit was quantified as 1 minus the absolute difference between the observed and expected allele frequencies

for years where allele frequencies were measured). This was done assuming complete dominance (h=1), assuming codominance (h=0.5), and with a random dominance coefficient to identify the combination of selection and dominance coefficients that maximised the fit to the data (Supplementary Table S6; supplementary Figure S9). The model and optimisation algorithm are available at the Github repository listed under 'Data Availability'.

We emphasise that our aim was to determine whether our estimate of the selection coefficient was reasonable, *i.e.* whether it was in the ballpark of values that could explain independently observed allele frequency estimates. This approach is not appropriate for determining a precise estimate of the selection coefficient; four important caveats should be considered. First, we estimate a single selection coefficient over hundreds of generations. Realistically, the strength of selection imposed by pesticide exposure will vary substantially over space and time. Since the effect of selection on allele frequency depends not only on its strength but on the standing frequency, our estimate will not necessarily reflect the average strength of selection over time. Second, we compare our retrodiction with empirical data that is itself only a course estimate of the actual allele frequency through time; Taylor *et al.* sampled between 22 and 52 chromosomes, so sampling error substantially affects the observed allele frequency. Third, only three timepoints are used, and since we model Δp from q_{1998} (to avoid any assumptions about when and how the allele first arose), we are only able to compare our estimates to two timepoints. The timepoints are, however, conveniently spaced for distinguishing between the distinct allele trajectories of dominant versus recessive mutations.

The fourth and most important caveat is that our comparison of $\hat{s}$ and $\hat{s}_{fit}$ is qualitative as substantially different methods are used for these estimates. Durrett and Schweinsberg's (2004) approximation of α does not take dominance into consideration. Maynard Smith and Haigh (1974) show that dominance affects the extent to which genetic diversity is reduced at neighbouring sites through hitch-hiking, but that this effect is still highly localised. Therefore, the estimate of *s* based on SweepFinder's α may be an over-estimation if the advantageous derived *CYP333B3* allele is completely dominant, and is not directly comparable with estimates of *s* that assume the advantageous allele is dominant. Allele frequencies modelled in Figure 8B are based on $\hat{s}$ yet assume complete dominance of the derived allele, so they should not strictly be interpreted as expected allele frequencies. Supplementary Figure S8 shows expected allele frequencies under a consistent assumption of co-dominance and produces qualitatively similar results. Although assuming co-dominance may allow for a more consistent comparison, this assumption may be biologically unrealistic: many cases of cytochrome P450-mediated pesticide resistance involve dominance coefficients >0.5, and we expect a gain-of function mutation allowing xenobiotic metabolism to act in a dominant manner (see Discussion; Cariño et al. 1994; Han et al. 2014; Sayyed et al. 2008; Achaleke & Brévault 2010; Heckel et al. 1998). Moreover, $\hat{s}_{fit}$ was a substantially worse fit to the empirical data when codominance was assumed – the observed rate of change early in time was far more consistent with selection on a dominant mutation (Supplementary Table S6). In this study at least, the difficult decision between directly comparable estimates and biologically realistic assumptions is resolved by the fact that $\hat{s}_{fit}$ does not vary by more than ~0.05 for dominance coefficients between 0.5 and 1 (see supplementary Figure S); differences between s estimates based on historical allele frequencies and estimates of s based on sweep parameters are similar irrespective of whether we assume the mutation is codominant, dominant, or anywhere in between. Therefore, despite these caveats, our estimate $\hat{s}$ is entirely consistent with independently measured shifts in allele frequency over time.

Testing for cryptic signatures of adaptive introgression at *CYP333B3*

We used three metrics to confirm that the sweep at *CYP333B3* was due to recent selection within *H. zea* and not the result of adaptive introgression from *H. armigera*. Adaptive introgression should result in decreased genetic differentiation from the donor species and an increased time to the most recent common ancestor (TMRCA) between homologous alleles at sites abutting the selected locus, and a decrease at the locus itself (Setter et al. 2020). By contrast, recent selection should result in an increase in genetic differentiation compared to samples collected earlier in time and a decreased TMRCA at the locus and linked sites. Therefore, we calculated F_{ST} between *H.* zea samples from 2002 vs 2019, as well as F_{ST} between *H. armigera* and *H. zea* sampled in 2019. This was done in both 20kbp and 100kbp windows. Next, we used Gamma-SMC (Schweiger & Durbin 2023) to estimate the TMRCA between homologous alleles at each polymorphic site on chromosome 13 within individuals sampled in 2019, using the average chromosome 13 scaled recombination rate estimate $\bar{\rho} = 0.217$ as a prior. This was only done within individuals (as opposed to between alleles of different individuals) so that bias resulting from potential phasing errors could be ruled out. All such analyses were carried out using call set 1.

Citations

Abd-Elghafar SF, Knowles CO, Wall ML. 1993. Pyrethroid Resistance in Two Field Strains of Helicoverpa zea (Lepidoptera: Noctuidae). J Econ Entomol. 86:1651–1655. doi: 10.1093/JEE/86.6.1651.

Achaleke J, Brévault T. 2010. Inheritance and stability of pyrethroid resistance in the cotton bollworm Helicoverpa armigera (Lepidoptera: Noctuidae) in Central Africa. Pest Manag Sci. 66:137–141. doi: 10.1002/PS.1843.

Ali MI, Luttrell RG, Young SY. 2006. Susceptibilities of Helicoverpa zea and Heliothis virescens (Lepidoptera: Noctuidae) Populations to Cry1Ac Insecticidal Protein. J. Econ. Entomol. 99:164–175.

Amezian D, Nauen R, Le Goff G. 2021. Comparative analysis of the detoxification gene inventory of four major Spodoptera pest species in response to xenobiotics. Insect Biochem Mol Biol. 138:103646. doi: 10.1016/J.IBMB.2021.103646.

Anderson CJ et al. 2018. Hybridization and gene flow in the mega-pest lineage of moth, Helicoverpa. Proceedings of the National Academy of Sciences. 201718831.

Arends B et al. 2021. Effectiveness of the natural resistance management refuge for Bt-cotton is dominated by local abundance of soybean and maize. Scientific Reports 2021 11:1. 11:1–9. doi: 10.1038/s41598-021-97123-8.

Arends BR et al. 2022. Helicoverpa zea (Lepidoptera: Noctuidae) feeding incidence and survival on Bt maize in relation to maize in the landscape. Pest Manag Sci. 78:2309–2315. doi: 10.1002/PS.6855.

van der Auwera GA et al. 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr Protoc Bioinformatics. 43:11.10.1-11.10.33. doi: 10.1002/0471250953.BI1110S43.

Bahia G. 2021. Examining Neonicotinoid Resistance in the Honeybee. York University: Ontario, Toronto.

Benowitz KM et al. 2022. Novel genetic basis of resistance to Bt toxin Cry1Ac in Helicoverpa zea. Genetics. 221. doi: 10.1093/GENETICS/IYAC037.

Bilbo TR, Reay-Jones FPF, Reisig DD, Greene JK. 2019. Susceptibility of Corn Earworm (Lepidoptera: Noctuidae) to Cry1A.105 and Cry2Ab2 in North and South Carolina. J Econ Entomol. 112:1845–1857. doi: 10.1093/JEE/TOZ062.

1188 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
1189 Bioinformatics. 30:2114–2120. doi: 10.1093/BIOINFORMATICS/BTU170.
1190 Bras A, Roy A, Heckel DG, Anderson P, Karlsson Green K. 2022. Pesticide resistance in arthropods:
1191 Ecology matters too. Ecol Lett. 25:1746–1759. doi: 10.1111/ELE.14030.
1192 Broad Institute. 2023. Picard toolkit. GitHub repository. http://broadinstitute.github.io/picard/
1193 (Accessed September 15, 2023).
1194 Browning BL, Browning SR. 2016. Genotype Imputation with Millions of Reference Samples. The
1195 American Journal of Human Genetics. 98:116–126. doi: 10.1016/J.AJHG.2015.11.020.
1196 Browning SR, Browning BL. 2007. Rapid and Accurate Haplotype Phasing and Missing-Data Inference
1197 for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. The
1198 American Journal of Human Genetics. 81:1084–1097. doi: 10.1086/521987.
1199 Brun-Barale A et al. 2010. Multiple P450 genes overexpressed in deltamethrin-resistant strains of
1200 Helicoverpa armigera. Pest Manag Sci. 66:900–909. doi: 10.1002/PS.1960.
1201 Cammarata-Mouchtouris A et al. 2020. Hyd ubiquitinates the NF-κB co-factor Akirin to operate an
1202 effective immune response in Drosophila. PLoS Pathog. 16:e1008458. doi:
1203 10.1371/JOURNAL.PPAT.1008458.
1204 Cariño FA, Koener JF, Plapp FW, Feyereisen R. 1994. Constitutive overexpression of the cytochrome
1205 P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect
1206 Biochem Mol Biol. 24:411–418. doi: 10.1016/0965-1748(94)90034-5.
1207 Chan AH, Jenkins PA, Song YS. 2012. Genome-Wide Fine-Scale Recombination Rate Variation in
1208 Drosophila melanogaster. PLoS Genet. 8:e1003090. doi: 10.1371/JOURNAL.PGEN.1003090.
1209 Chang CC et al. 2015. Second-generation PLINK: Rising to the challenge of larger and richer datasets.
1210 Gigascience. 4:7. doi: 10.1186/S13742-015-0047-8/2707533.
1211 Chen MZ et al. 2021. Migration trajectories of the diamondback moth Plutella xylostella in China
1212 inferred from population genomic variation. Pest Manag Sci. 77:1683–1693. doi:
1213 10.1002/PS.6188.
1214 Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics.
1215 34:i884–i890. doi: 10.1093/BIOINFORMATICS/BTY560.
1216 Chown SL et al. 2015. Biological invasions, climate change and genomics. Evol Appl. 8:23–46. doi:
1217 10.1111/EVA.12234.
1218 Clark RD et al. 2023. The practice and promise of temporal genomics for measuring evolutionary
1219 responses to global change. Mol Ecol Resour. doi: 10.1111/1755-0998.13789.
1220 Cook D, Threet M. 2019. Cotton Crop Losses 2019. Mississippi Agricultural and Forestry Experiment
1221 Station. https://www.biochemistry.msstate.edu/resources/2019loss.php (Accessed April 17,
1222 2023).
1223 Cunningham JP, Zalucki MP. 2014. Understanding Heliothine (Lepidoptera: Heliothinae) Pests: What is
1224 a Host Plant? J Econ Entomol. 107:881–896. doi: 10.1603/EC14036.
1225 Czech L, Huerta-Cepas J, Stamatakis A. 2017. A Critical Review on the Use of Support Values in Tree
1226 Viewers and Bioinformatics Toolkits. Mol Biol Evol. 34:1535–1542. doi:
1227 10.1093/MOLBEV/MSX055.
1228 Danecek P et al. 2011. The variant call format and VCFtools. Bioinformatics. 27:2156–2158. doi:
1229 10.1093/BIOINFORMATICS/BTR330.
1230 DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R. 2016. SweepFinder2: increased sensitivity,
1231 robustness and flexibility. Bioinformatics. 32:1895–1897.
1232 Deutsch CA et al. 2018. Increase in crop losses to insect pests in a warming climate. Science (1979).
1233 361:916–919. doi: 10.1126/SCIENCE.AAT3466.
1234 Dively GP, Venugopal PD, Finkenbinder C. 2016. Field-Evolved Resistance in Corn Earworm to Cry
1235 Proteins Expressed by Transgenic Sweet Corn. PLoS One. 11:e0169115. doi:
1236 10.1371/JOURNAL.PONE.0169115.

1237 Durigan MR et al. 2017. High frequency of CYP337B3 gene associated with control failures of
1238 Helicoverpa armigera with pyrethroid insecticides in Brazil. Pestic Biochem Physiol. 143:73–
1239 80. doi: 10.1016/J.PESTBP.2017.09.005.
1240 Durrett R, Schweinsberg J. 2004. Approximating selective sweeps. Theor Popul Biol. 66:129–138.
1241 Fitt GP. 1989. The Ecology of Heliothis Species in Relation to Agroecosystems. Annu Rev Entomol.
1242 34:17–52.
1243 Flack JE, Mieszczanek J, Novcic N, Bienz M. 2017. Wnt-Dependent Inactivation of the Groucho/TLE Co-
1244 repressor by the HECT E3 Ubiquitin Ligase Hyd/UBR5. Mol Cell. 67:181-193.e5. doi:
1245 10.1016/J.MOLCEL.2017.06.009.
1246 Gordon KHJ, Tay WT, Collinge D, Williams A, Batterham P. 2010. Genetics and molecular biology of
1247 the major crop pest genus helicoverpa. In: Molecular Biology and Genetics of the
1248 Lepidoptera. CRC Press pp. 219–233.
1249 Gould F, Brown ZS, Kuzma J. 2018. Wicked evolution: Can we address the sociobiological dilemma of
1250 pesticide resistance? Science (1979). 360:728–732. doi: 10.1126/SCIENCE.AAR3780.
1251 Haag-Liautard C et al. 2007. Direct estimation of per nucleotide and genomic deleterious mutation
1252 rates in Drosophila. Nature 2006 445:7123. 445:82–85. doi: 10.1038/nature05388.
1253 Hamelin RC, Roe AD. 2020. Genomic biosurveillance of forest invasive alien enemies: A story written
1254 in code. Evol Appl. 13:95–115. doi: 10.1111/EVA.12853.
1255 Han H, Yu Q, Vila-Aiub M, Powles SB. 2014. Genetic inheritance of cytochrome P450-mediated
1256 metabolic resistance to chlorsulfuron in a multiple herbicide resistant Lolium rigidum
1257 population. Crop Protection. 65:57–63. doi: 10.1016/J.CROPRO.2014.06.026.
1258 Han Yangchun. 2014. Metabolic mechanism, QTL mapping and transcriptomic analysis of resistance to
1259 fenvalerate in helicoverpa armigera. Nanjing Agricultural University: Nanjing.
1260 Hardwick D. 1965. The corn earworm complex. Memoirs of the Entomological Society of Canada.
1261 97:5–247.
1262 Hawkins NJ, Bass C, Dixon A, Neve P. 2019. The evolutionary origins of pesticide resistance. Biological
1263 Reviews. 94:135–155. doi: 10.1111/BRV.12440.
1264 Heckel DG, Gahan LJ, Daly JC, Trowell S. 1998. A genomic approach to understanding Heliothis and
1265 Helicoverpa resistance to chemical and biological insecticides. Philos Trans R Soc Lond B Biol
1266 Sci. 353:1713–1722.
1267 Hessenauer P et al. 2020. Hybridization and introgression drive genome evolution of Dutch elm
1268 disease pathogens. Nature Ecology & Evolution 2020 4:4. 4:626–638. doi: 10.1038/s41559-
1269 020-1133-6.
1270 Huang F et al. 2023. An Extended Investigation of Unexpected Helicoverpa zea (Boddie) Survival and
1271 Ear Injury on a Transgenic Maize Hybrid Expressing Cry1A/Cry2A/Vip3A Toxins. Toxins 2023,
1272 Vol. 15, Page 474. 15:474. doi: 10.3390/TOXINS15070474.
1273 Ibrahim SS et al. 2015. Allelic Variation of Cytochrome P450s Drives Resistance to Bednet Insecticides
1274 in a Major Malaria Vector. PLoS Genet. 11:e1005618. doi: 10.1371/JOURNAL.PGEN.1005618.
1275 Ivey V, Hillier NK. 2023. Hybridization in heliothine moths: impacts on reproduction, pheromone
1276 communication, and pest management. Front Ecol Evol. 11:1208079. doi:
1277 10.3389/FEVO.2023.1208079/BIBTEX.
1278 Jacobson A et al. 2009. Resistance to Pyrethroid Insecticides in Helicoverpa zea (Lepidoptera:
1279 Noctuidae) in Indiana and Illinois. J Econ Entomol. 102:2289–2295. doi:
1280 10.1603/029.102.0634.
1281 Jia ZQ et al. 2020. Identification of transcriptome and fluralaner responsive genes in the common
1282 cutworm Spodoptera litura Fabricius, based on RNA-seq. BMC Genomics. 21:1–11. doi:
1283 10.1186/S12864-020-6533-0/TABLES/1.
1284 Jin M et al. 2023. Adaptive evolution to the natural and anthropogenic environment in a global
1285 invasive crop pest, the cotton bollworm. The Innovation. 4:100454. doi:
1286 10.1016/J.XINN.2023.100454.

1287 Jones CM, Parry H, Tay WT, Reynolds DR, Chapman JW. 2018. Movement Ecology of Pest Helicoverpa:
1288 Implications for Ongoing Spread. doi: 10.1146/annurev-ento-011118.
1289 Joußen N et al. 2012. Resistance of Australian Helicoverpa armigera to fenvalerate is due to the
1290 chimeric P450 enzyme CYP337B3. Proceedings of the National Academy of Sciences.
1291 201202047.
1292 Karlsson Green K, Stenberg JA, Lankinen Å. 2020. Making sense of Integrated Pest Management (IPM)
1293 in the light of evolution. Evol Appl. 13:1791–1805. doi: 10.1111/EVA.13067.
1294 Karney CFF. 2013. Algorithms for geodesics. 87:43–55. doi: 10.1007/s00190-012-0578-z.
1295 Kim Y, Stephan W. 2000. Joint Effects of Genetic Hitchhiking and Background Selection on Neutral
1296 Variation. Genetics. 155:1415–1427. doi: 10.1093/GENETICS/155.3.1415.
1297 Korunes KL, Samuk K. 2021. pixy: Unbiased estimation of nucleotide diversity and divergence in the
1298 presence of missing data. Mol Ecol Resour. 21:1359–1368. doi: 10.1111/1755-0998.13326.
1299 Kreiner JM et al. 2019. Multiple modes of convergent adaptation in the spread of glyphosate-resistant
1300 Amaranthus tuberculatus. Proc Natl Acad Sci U S A. 116:21076–21084. doi:
1301 10.1073/PNAS.1900870116/SUPPL_FILE/PNAS.1900870116.SAPP.PDF.
1302 Kriticos DJ et al. 2015. The potential distribution of invading Helicoverpa armigera in North America: is
1303 it just a matter of time? PLoS One. 10:e0119618.
1304 de la Paz Celorio-Mancera M, Ahn SJ, Vogel H, Heckel DG. 2011. Transcriptional responses underlying
1305 the hormetic and detrimental effects of the plant secondary metabolite gossypol on the
1306 generalist herbivore Helicoverpa armigera. BMC Genomics. 12:575. doi: 10.1186/1471-2164-
1307 12-575.
1308 Lawton D et al. 2022. Pest population dynamics are related to a continental overwintering gradient.
1309 Proc Natl Acad Sci U S A. 119:e2203230119. doi:
1310 10.1073/PNAS.2203230119/SUPPL_FILE/PNAS.2203230119.SM01.MP4.
1311 Li H et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.
1312 doi: 10.1093/BIOINFORMATICS/BTP352.
1313 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform.
1314 Bioinformatics. 25:1754–1760. doi: 10.1093/BIOINFORMATICS/BTP324.
1315 Li YF, Costello JC, Holloway AK, Hahn MW. 2008. “Reverse ecology” and the power of population
1316 genomics. Evolution. 62:2984–2994.
1317 Lilley CE et al. 2010. A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and
1318 DNA damage responses. EMBO J. 29:943–955. doi: 10.1038/EMBOJ.2009.400.
1319 Love RR, Sikder JR, Vivero RJ, Matute DR, Schrider DR. 2023. Strong positive selection in Aedes aegypti
1320 and the rapid evolution of insecticide resistance Nielsen, R, editor. Mol Biol Evol. doi:
1321 10.1093/MOLBEV/MSAD072.
1322 Luke SH et al. 2023. Grand challenges in entomology: Priorities for action in the coming decades.
1323 Insect Conserv Divers. 16:173–189. doi: 10.1111/ICAD.12637.
1324 Luttrell RG, Jackson RE. GM Crops & Food Biotechnology in Agriculture and the Food Chain
1325 Helicoverpa zea and Bt Cotton in the United States. 3:213–227. doi: 10.4161/gmcr.20742.
1326 Madeira F et al. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic
1327 Acids Res. 50:W276. doi: 10.1093/NAR/GKAC240.
1328 Mallet J. 1989. The evolution of insecticide resistance: Have the insects won? Trends Ecol Evol. 4:336–
1329 340. doi: 10.1016/0169-5347(89)90088-8.
1330 Margosian ML, Garrett KA, Hutchinson JMS, With KA. 2009. Connectivity of the American Agricultural
1331 Landscape: Assessing the National Risk of Crop Pest and Disease Spread. Bioscience. 59:141–
1332 151. doi: 10.1525/BIO.2009.59.2.7.
1333 Marques DA, Meier JI, Seehausen O. 2019. A Combinatorial View on Speciation and Adaptive
1334 Radiation. Trends Ecol Evol. 34:531–544. doi:
1335 10.1016/J.TREE.2019.02.008/ATTACHMENT/7187B819-180D-4801-8F86-
1336 52726E4C683F/MMC1.DOCX.

1337 Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the Use of ABBA–BABA Statistics to Locate
1338 Introgressed Loci. Mol Biol Evol. 32:244–257. doi: 10.1093/MOLBEV/MSU269.
1339 Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to
1340 introgression across butterfly genomes. PLoS Biol. 17:e2006288. doi:
1341 10.1371/JOURNAL.PBIO.2006288.
1342 Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. Genet Res. 23:23–35.
1343 Mazzi D, Dorn S. 2012. Movement of insect pests in agricultural landscapes. Annals of Applied Biology.
1344 160:97–113. doi: 10.1111/J.1744-7348.2012.00533.X.
1345 McDonald BA, Linde C. 2003. Pathogen population genetics, evolutionary potential, and durable
1346 resistance. Annu Rev Phytopathol. 40:349–379.
1347 Mehta S V., Haight RG, Homans FR, Polasky S, Venette RC. 2007. Optimal detection and control
1348 strategies for invasive species management. Ecological Economics. 61:237–245. doi:
1349 10.1016/J.ECOLECON.2006.10.024.
1350 Mölder F et al. 2021. Sustainable data analysis with Snakemake. F1000Research 2021 10:33. 10:33.
1351 doi: 10.12688/f1000research.29032.1.
1352 Morey AC. 2010. Corn earworm (Helicoverpa zea Boddie), cold hardiness, and climate change:
1353 Implications for future distributions and IPM. University of Minnesota.
1354 Van Munster M et al. 2007. Altered gene expression in Choristoneura fumiferana and Manduca sexta
1355 in response to sublethal intoxication by Bacillus thuringiensis Cry1Ab toxin. Insect Mol Biol.
1356 16:25–35. doi: 10.1111/j.1365-2583.2006.00692.x.
1357 Nauen R, Bass C, Feyereisen R, Vontas J. 2022. The Role of Cytochrome P450s in Insect Toxicology and
1358 Resistance. Annual Review of Entomology . 67:105–124.
1359 Neafsey DE, Taylor AR, MacInnis BL. 2021. Advances and opportunities in malaria population
1360 genomics. Nature Reviews Genetics 2021 22:8. 22:502–517. doi: 10.1038/s41576-021-00349-
1361 5.
1362 Neph S et al. 2012. BEDOPS: high-performance genomic feature operations. Bioinformatics. 28:1919–
1363 1920. doi: 10.1093/BIOINFORMATICS/BTS277.
1364 Ni R et al. 2023. Absence of known knockdown resistance mutations but fixation of CYP337B3 was
1365 detected in field populations of Helicoverpa armigera across China. Pestic Biochem Physiol.
1366 195:105542. doi: 10.1016/J.PESTBP.2023.105542.
1367 Niedermann S, Crameri S, Jensen B, Tanner M. 2017. Integrated pest and resistance management of
1368 Helicoverpa zea (Lepidoptera: Noctuidae) with a Helicoverpa armigera nucleopolyhedrovirus.
1369 In: Proceedings of hte IOBC/WPRS Working Group 'Microbial and Nematode Control of
1370 Invertebrate Pests'.Vol. 129 International Organization for Biological and Integrated Control
1371 of Noxious Animals and Plants (OIBC/OILB), West Palaearctic Regional Section (WPRS/SROP):
1372 Tbilisi pp. 189–193.
1373 Nielsen R et al. 2005. Genomic scans for selective sweeps using SNP data. Genome Res. 15:1566–
1374 1575. doi: 10.1101/GR.4252305.
1375 North HL, McGaughran A, Jiggins CD. 2021. Insights into invasive species from whole-genome
1376 resequencing. Mol Ecol. 30:6289–6308. doi: 10.1111/MEC.15999.
1377 Obbard DJ, Jiggins FM, Bradshaw NJ, Little TJ. 2011. Recent and Recurrent Selective Sweeps of the
1378 Antiviral RNAi Gene Argonaute-2 in Three Species of Drosophila. Mol Biol Evol. 28:1043–1056.
1379 doi: 10.1093/MOLBEV/MSQ280.
1380 Olivi BM et al. 2019. Impact of Simulated Corn Earworm (Lepidoptera: Noctuidae) Kernel Feeding on
1381 Field Corn Yield. J Econ Entomol. 112:2193–2198. doi: 10.1093/JEE/TOZ119.
1382 Olmstead DL, Nault BA, Shelton AM. 2016. Biology, Ecology, and Evolving Management of Helicoverpa
1383 zea (Lepidoptera: Noctuidae) in Sweet Corn in the United States. J Econ Entomol. 109:1667–
1384 1676. doi: 10.1093/JEE/TOW125.
1385 Olmstead DL, Nault BA, Shelton AM. Field and Forage Crops Biology, Ecology, and Evolving
1386 Management of Helicoverpa zea (Lepidoptera: Noctuidae) in Sweet Corn in the United States.
1387 doi: 10.1093/jee/tow125.

Oziolor EM et al. 2019. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. Science (1979). 364:455–457. doi: 10.1126/SCIENCE.AAV4155/SUPPL_FILE/AAV4155-OZIOLOR-SM.PDF.

Paini DR et al. 2016. Global threat to agriculture from invasive species. Proc Natl Acad Sci U S A. 113:7575–7579. doi: 10.1073/PNAS.1602205113.

Parajulee MN, Rummel DR, Arnold MD, Carroll SC. 2004. Long-Term Seasonal Abundance Patterns of Helicoverpa zea and Heliothis virescens (Lepidoptera: Noctuidae) in the Texas High Plains. J Econ Entomol. 97:668–677.

Paris V et al. 2022. Urban population structure and dispersal of an Australian mosquito (Aedes notoscriptus) involved in disease transmission. Heredity 2022 130:2. 130:99–108. doi: 10.1038/s41437-022-00584-4.

Pearce SL et al. 2017. 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.

Pélissié B, Crossley MS, Cohen ZP, Schoville SD. 2018. Rapid evolution in insect pests: the importance of space and time in population genomics studies. Curr Opin Insect Sci. 26:8–16. doi: 10.1016/J.COIS.2017.12.008.

Peng X, Wang S, Huang L, Su S, Chen M. 2021. Characterization of Rhopalosiphum padi takeout-like genes and their role in insecticide susceptibility. Pestic Biochem Physiol. 171:104725. doi: 10.1016/J.PESTBP.2020.104725.

Perera OP, Fescemyer HW, Fleischer SJ, Abel CA. 2020. Temporal Variation in Genetic Composition of Migratory Helicoverpa Zea in Peripheral Populations. Insects 2020, Vol. 11, Page 463. 11:463. doi: 10.3390/INSECTS11080463.

Petit N, Barbadilla A. 2009. Selection efficiency and effective population size in Drosophila species. J Evol Biol. 22:515–526. doi: 10.1111/J.1420-9101.2008.01672.X.

Pool JE, Nielsen R. 2009. Inference of Historical Changes in Migration Rate From the Lengths of Migrant Tracts. Genetics. 181:711–719. doi: 10.1534/GENETICS.108.098095.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842.

Rasool A et al. 2014. An independent occurrence of the chimeric P450 enzyme CYP337B3 of Helicoverpa armigera confers cypermethrin resistance in Pakistan. Insect Biochem Mol Biol. 53:54–65. doi: 10.1016/J.IBMB.2014.07.006.

Reisig DD et al. 2022. Best Management Practices to Delay the Evolution of Bt Resistance in Lepidopteran Pests Without High Susceptibility to Bt Toxins in North America. J Econ Entomol. 115:10–25. doi: 10.1093/JEE/TOAB247.

Reisig DD et al. 2018. Long-Term Empirical and Observational Evidence of Practical Helicoverpa zea Resistance to Cotton With Pyramided Bt Toxins. J Econ Entomol. 111:1824–1833. doi: 10.1093/JEE/TOY106.

Reisig DD, Kurtz R. 2018. Bt Resistance Implications for Helicoverpa zea (Lepidoptera: Noctuidae) Insecticide Resistance Management in the United States. Environ Entomol. 47:1357–1364. doi: 10.1093/EE/NVY142.

Rousset F. 1997. Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. Genetics. 145:1219–1228. doi: 10.1093/GENETICS/145.4.1219.

Sayyed AH, Ahmad M, Saleem MA. 2008. Cross-Resistance and Genetics of Resistance to Indoxacarb in Spodoptera litura (Lepidoptera: Noctuidae). J Econ Entomol. 101:472–479.

Schweiger R, Durbin R. 2023. Ultra-fast genome-wide inference of pairwise coalescence times. bioRxiv. 2023.01.06.522935. doi: 10.1101/2023.01.06.522935.

Setter D et al. 2020. VolcanoFinder: Genomic scans for adaptive introgression. PLoS Genet. 16:e1008867. doi: 10.1371/JOURNAL.PGEN.1008867.

Seymour M et al. 2016. Peripheral genetic structure of Helicoverpa zea indicates asymmetrical panmixia. Ecol Evol. 6:3198–3207. doi: 10.1002/ECE3.2106.

1439 Sherpa S, Després L. 2021. The evolutionary dynamics of biological invasions: A multi-approach
1440 perspective. Evol Appl. 14:1463–1484. doi: 10.1111/EVA.13215.
1441 Shi Y et al. 2022. Involvement of CYP2 and mitochondrial clan P450s of Helicoverpa armigera in
1442 xenobiotic metabolism. Insect Biochem Mol Biol. 140:103696. doi:
1443 10.1016/J.IBMB.2021.103696.
1444 Da Silva IHS et al. 2018. Identification of midgut membrane proteins from different instars of
1445 Helicoverpa armigera (Lepidoptera: Noctuidae) that bind to Cry1Ac toxin. PLoS One.
1446 13:e0207789. doi: 10.1371/JOURNAL.PONE.0207789.
1447 Song Y et al. 2011. Adaptive Introgression of Anticoagulant Rodent Poison Resistance by Hybridization
1448 between Old World Mice. Current Biology. 21:1296–1301. doi: 10.1016/j.cub.2011.06.043.
1449 Specht A et al. 2021. Helicoverpa armigera (Hübner) (Lepidoptera: Noctuidae) in Brazil: the Big
1450 Outbreak Monitored by Light Traps. Neotrop Entomol. 50:53–67. doi: 10.1007/S13744-020-
1451 00836-0.
1452 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
1453 phylogenies. Bioinformatics. 30:1312–1313. doi: 10.1093/BIOINFORMATICS/BTU033.
1454 Tay WT et al. 2013. A Brave New World for an Old World Pest: Helicoverpa armigera (Lepidoptera:
1455 Noctuidae) in Brazil. PLoS One. 8:e80134. doi: 10.1371/JOURNAL.PONE.0080134.
1456 Tay WT et al. 2022. Global population genomic signature of Spodoptera frugiperda (fall armyworm)
1457 supports complex introduction events across the Old World. Communications Biology 2022
1458 5:1. 5:1–15. doi: 10.1038/s42003-022-03230-1.
1459 Tay WT, Gordon KHJ. 2019. Going global - genomic insights into insect invasions. Curr Opin Insect Sci.
1460 31:123–130. doi: 10.1016/J.COIS.2018.12.002.
1461 Taylor KL, Hamby KA, DeYonke AM, Gould F, Fritz ML. 2021. Genome evolution in an agricultural pest
1462 following adoption of transgenic crops. Proc Natl Acad Sci U S A. 118:e2020853118. doi:
1463 10.1073/PNAS.2020853118/SUPPL_FILE/PNAS.2020853118.SD06.XLSX.
1464 Tchouakui M et al. 2021. Combined over-expression of two cytochrome P450 genes exacerbates the
1465 fitness cost of pyrethroid resistance in the major African malaria vector Anopheles funestus.
1466 Pestic Biochem Physiol. 173:104772. doi: 10.1016/J.PESTBP.2021.104772.
1467 Tembrock LR, Timm AE, Zink FA, Gilligan TM. 2019. Phylogeography of the Recent Expansion of
1468 Helicoverpa armigera (Lepidoptera: Noctuidae) in South America and the Caribbean Basin.
1469 Ann Entomol Soc Am. 112:388–401. doi: 10.1093/AESA/SAZ019.
1470 Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010. A Generation Time Effect on the Rate of Molecular
1471 Evolution in Invertebrates. Mol Biol Evol. 27:1173–1180. doi: 10.1093/MOLBEV/MSQ009.
1472 Toé KH, N'Falé S, Dabiré RK, Ranson H, Jones CM. 2015. The recent escalation in strength of
1473 pyrethroid resistance in Anopheles coluzzi in West Africa is linked to increased expression of
1474 multiple gene families. BMC Genomics. 16:1–11. doi: 10.1186/S12864-015-1342-
1475 6/FIGURES/5.
1476 USDA APHIS. Old World Bollworm: where's the threat?
1477 https://www.aphis.usda.gov/aphis/resources/pests-diseases/hungry-pests/the-threat/old-
1478 world-bollworm/owb (Accessed September 7, 2023).
1479 Valencia-Montoya WA et al. 2020. Adaptive Introgression across Semipermeable Species Boundaries
1480 between Local Helicoverpa zea and Invasive Helicoverpa armigera Moths. Mol Biol Evol.
1481 37:2568–2583. doi: 10.1093/MOLBEV/MSAA108.
1482 Walsh TK et al. 2022. Determinants of Insecticide Resistance Evolution: Comparative Analysis Among
1483 Heliothines. Annual Review of Entomology . 67:387–406. doi: 10.1146/ANNUREV-ENTO-
1484 080421-071655.
1485 Wang K et al. 2023. Atrazine exposure can dysregulate the immune system and increase the
1486 susceptibility against pathogens in honeybees in a dose-dependent manner. J Hazard Mater.
1487 452:131179. doi: 10.1016/J.JHAZMAT.2023.131179.
1488 Westbrook JK. 2008. Noctuid migration in Texas within the nocturnal aeroecological boundary layer.
1489 Integr Comp Biol. 48:99–106. doi: 10.1093/ICB/ICN040.

Xu S et al. 2022. Ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. iMeta. 1:e56. doi: 10.1002/IMT2.56.

Yang Y, Xu H, Lu Y, Wang C, Lu Z. 2018. Midgut transcriptomal response of the rice leaffolder, Cnaphalocrocis medinalis (Guenée) to Cry1C toxin. PLoS One. 13:e0191686. doi: 10.1371/JOURNAL.PONE.0191686.

Yu W et al. 2021. Extended investigation of field-evolved resistance of the corn earworm Helicoverpa zea (Lepidoptera: Noctuidae) to Bacillus thuringiensis Cry1A.105 and Cry2Ab2 proteins in the southeastern United States. J Invertebr Pathol. 183:107560. doi: 10.1016/J.JIP.2021.107560.

Zhang J et al. 2022. Population genomics provides insights into lineage divergence and local adaptation within the cotton bollworm. Mol Ecol. 22:1875–1891. doi: 10.1111/1755-0998.13581.

Data availability statement

All raw .fastq files generated here are available on the Short Read Archive: biosample accessions SAMN27502736-SAMN27502972, along with call sets as .vcf files. All custom scripts, processed data and scripts required to reproduce figures, and supplementary tables, are available at https://github.com/hlnorth/north_american_helicoverpa_zea

Funding statement

This work was made possible through funding from a bilateral BBSRC-FAPESP grant awarded to CJ (project reference: BB/V001329/1). This work was funded in part by a USDA-APHIS Cooperative Agreement (award no. AP18PPQS&T00C204) to GS. HLN was supported by funding from the Department of Zoology, University of Cambridge.

Conflict of interest disclosure

The authors declare no conflict of interest.

Author contributions

HLN designed and implemented all analyses in all call sets, apart from variant calling in call set 2, under the supervision of CJ, with feedback and advice from ZF and GS. HLN wrote the manuscript text and prepared figures under the supervision of CJ with feedback from all co-authors. ZF conceived the field sampling design, collected samples, performed variant calling for call set 2, and conducted preliminary analyses not included in this publication under the supervision of GS. Other authors were integral to facilitating sample collection. We are grateful to Joana Meier and all members of the Insect Evolution and Genomics Group at Cambridge for feedback and advice on the analyses. Kyle Benowitz kindly provided the susceptible and resistant laboratory strain sequences of *Kinesin-12*, which allowed us to verify genotype calls.