Corresponding Author:

Dr. Megan L. Fritz
Department of Entomology
University of Maryland
4291 Field House Dr.
Plant Sciences Bldg. Rm. 4112
College Park, MD 20742
mfritz13@umd.edu

Genome evolution in an agricultural pest following adoption of transgenic crops

Megan L. Fritz[1*], Kelly A. Hamby[1], Katherine Taylor[1], Alexandra M. DeYonke[2], and Fred Gould[2]

[1]Department of Entomology, University of Maryland, College Park, MD 20742 USA

[2]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27607 USA

* Corresponding author

Running Title: Resistance Evolution in Agricultural Ecosystems

Keywords

Helicoverpa zea, Bt resistance, temporal genomic change, polygenic selection

Abstract

Replacement of synthetic insecticides with transgenic crops for pest management has been both economically and environmentally beneficial. These benefits have often eroded as pests evolved resistance to the transgenic crops, but a broad understanding of the timing and complexity of the adaptive changes which lead to field-evolved resistance in pest species is lacking. Wild populations of *Helicoverpa zea*, a major lepidopteran crop pest and the target of transgenic Cry toxin-expressing cotton and corn, have recently evolved widespread, damaging levels of resistance. Here, we quantified patterns of genomic change in wild *H. zea* collected between 2002 and 2017 when adoption rates of Cry-expressing crops expanded in North America. Using a combination of genomic and genotypic approaches, we identified significant temporal changes in allele frequency throughout the genomes of field-collected *H. zea.* Many of these changes occurred concurrently with increasingly damaging levels of resistance to Cry toxins between 2012 and 2016, in a pattern consistent with polygenic selection. Surprisingly, none of the eleven previously described Cry resistance genes showed signatures of selection in wild *H. zea.* Furthermore, we observed evidence of a very strong selective sweep in one region of the *H. zea* genome, yet this strongest change was not additively associated with Cry resistance. This first, whole genome analysis of field-collected specimens to study evolution of Cry resistance demonstrates the potential and need for a more holistic approach to examining pest adaptation to changing agricultural practices.

43 **Introduction**

44 The study of rapid evolutionary responses to natural selection has gained momentum in the past
45 decade, and along with it recognition of the relevance of pest adaptation to control measures (Reznick
46 et al. 2019). Melander published the first observation of insecticide resistance evolution in 1914
47 (Melander 1914), and since then, pests and weeds have adapted to myriad chemical, biological, and
48 cultural control measures (Gould 1991, Palumbi 2001, Délye et al. 2013, Gould et al. 2018). While
49 there has been an emphasis on finding single genetic loci that confer resistance to pesticides,
50 recognition that surviving insecticide or herbicide applications could involve a suite of genomic
51 changes has increased (Baucom 2019). Yet a clear understanding of the timing and complexity of
52 changes associated with adaptation to agricultural ecosystems is lacking (Leon et al. 2020).

53 The bacterium, *Bacillus thuringiensis* (Bt), produces one or more crystalline (Cry) proteins that
54 are each toxic to a narrow range of insect pests when ingested. Genes encoding Cry toxins were
55 engineered into cultivated crops to reduce pest damage, producing so-called Bt crops, a major
56 innovation in agricultural pest management that has been adopted around the globe (NASEM 2016).
57 The economic and environmental benefits of this technology are significant: they can replace heavily
58 used synthetic insecticides, which have damaging off target effects on pollinators and natural enemies,
59 and have contributed to area wide pest suppression in non-Bt crops (Hutchison et al. 2010, NASEM
60 2016, Dively et al. 2018). Yet a number of pest species have evolved resistance to one or more Cry
61 toxins and there is concern about the durability of these transgenic crops (Tabashnik and Carrière
62 2017). Resistance management strategies have been implemented effectively for a few major pest
63 species with high susceptibility to the toxins (NASEM 2016). Unfortunately, these measures have
64 been inadequate for species where Cry toxins cause only moderate levels of mortality, providing
65 opportunities for the evolution of resistance (Tabashnik and Carrière 2017).

66 Discovery of resistance-conferring genes has been of long-standing interest, in part due to
67 expectations that this knowledge could lead to molecular diagnostic approaches that facilitate detection

68 of early resistance evolution in wild pest populations (US-EPA 1998, US-EPA 2018). Molecular
69 signals of resistance could then serve as triggers for regulatory actions that would minimize resistance
70 risk and preserve environmentally sustainable technologies, like Cry toxins (US-EPA 1998). Genes
71 discovered to have major impacts on Cry tolerance are being used to a limited extent for molecular
72 resistance monitoring (Morin et al. 2004, Gahan et al. 2007, Zhang et al. 2013, Jin et al. 2018). To
73 date, there are at least 11 such genes in the Lepidoptera (Wu 2014, Lin 2019), many of which were
74 discovered in laboratory-selected populations and from screening wild populations for single genes
75 with major effects on resistance (Morin et al. 2003, Gahan et al. 2001, Zhang et al. 2012, Jin et al.
76 2018). The question of whether molecular monitoring of these genes would effectively track resistance
77 evolution in wild populations remains unresolved.
78 *Helicoverpa zea* is a migratory pest species (Gould et al. 2002) which comprises a single
79 panmictic population in North America (Seymour et al. 2016). It has been controlled with Bt crops
80 since 1996, when corn and cotton expressing Cry1A became commercially available. In 2007, corn
81 and cotton cultivars producing both Cry1A and Cry2A were introduced. Since their
82 commercialization, North American adoption rates of Cry-expressing corn and cotton grew from 8 and
83 15% of corn and cotton acreage in 1997 to *ca.* 80% by 2012. This rapid expansion of Cry-expressing
84 crops placed significant selection pressure on their target pests, and widespread, damaging levels of
85 resistance to Cry-expressing crops by *H. zea* were documented by 2016 (Dively et al. 2016, Reisig et
86 al. 2018, Kaur et al. 2019, Yang et al. 2019). Existing Bt corn and cotton cultivars do not produce a
87 high dose of Cry protein relative to what is required to kill heterozygote *H. zea* larvae with resistance
88 alleles (Luttrell and Jackson 2012). This likely contributed to their recent resistant status (Tabashnik
89 and Carrière 2017). Under these conditions, resistance to Cry proteins was predicted to be polygenic
90 (Gould 1998, Tabashnik et al. 2004) and to arise from standing genetic variation (Roush and McKenzie
91 1987).

92 Between 2002 and 2017, we saved *ca.* 1000 adult moths in ethanol each year, and we used
93 samples from these collections to track genomic changes in *H. zea* over time. Instead of simply
94 monitoring for changes in previously identified candidate genes, we took a gene agnostic approach and
95 scanned entire genomes for signatures of selection. Using two sequencing approaches, double-digest
96 Restriction-site Associated DNA sequencing (ddRAD-seq) and whole genome resequencing, we
97 identified multiple genomic regions with significant changes in allele frequency during a 15 year time
98 period. Shifts in allele frequency were used to calculate the strength of selection imposed on wild *H.*
99 *zea* in agricultural ecosystems. We also examined whether any of the 11 previously identified
100 candidate Cry resistance genes showed evidence of strong shifts in allele frequency over time.
101 Surprisingly, none of these genes could be implicated in Cry-toxin resistance evolution for *H. zea*.
102 Instead, we found a number of novel candidate genes within regions of greater than expected allele
103 frequency change. Due to changes in synthetic insecticide use following the introduction of Bt crops,
104 we also examined whether the frequency of alleles of *H. zea* genes associated with detoxification of
105 synthetic insecticides or even plant defensive compounds had changed, as was observed in another
106 species where Bt crops replaced insecticidal control (Fritz et al. 2018).
107 This first whole genome analysis of field-collected specimens to study impacts of Bt crops
108 demonstrates the potential and need for a more holistic approach to examining pest adaptation to
109 changing agricultural practices. While monitoring previously identified candidate genes may prove
110 useful in some cases, our results indicate that such an approach is likely to miss critical evolutionary
111 events. Analysis of patterns of genomic changes in pest populations following the introduction of a
112 new control tactic or tool could prove useful, but requires follow-up experimental confirmation to
113 demonstrate clear, causal relationships among genes and adaptive phenotypes.
114
115

116 **Results**

117 *Strong temporal ddRAD-seq allele frequency shifts*

118 *H. zea* males were collected from five pheromone baited traps (Hartstack 1979) in Bossier
119 Parish, LA, and pooled by collection date in the months of May through October, 2002 through 2017.
120 Randomly sampled moths from a subset of years were selected for genetic analysis; 2002, 2007, 2012,
121 as well as in 2016, when strong phenotypic resistance to Cry1Ab and increasing resistance to Cry2Ab
122 was being documented (Dively et al. 2016, Kaur et al. 2019). Two hundred and sixty-five specimens
123 were prepared into ddRAD-seq libraries and sequenced according to Fritz *et al.* (2016, 2018).
124 Following filtering and alignment to the 341 Mb *H. zea* genome, we identified 14,398 single nucleotide
125 polymorphisms (SNPs) from 259 individuals with an average read count of 404,003 (Figure S1).
126 Nucleotide diversity (π), and heterozygosity (F) did not change between years (Table S2), and overall
127 genomic divergence was low according to Weir and Cockerham's F_{ST} ($F_{ST} < 0.004$; Table S3). This
128 low divergence was also supported by a sequential k-means analysis, which showed that a single
129 genetic cluster or population (k = 1) best described our ddRAD-seq genotyped *H. zea* (Figure S2).
130 Interestingly, genetic divergence that occurred between the 2012-2016 collection periods was at least 3
131 times higher ($F_{ST} = 0.0025$) than for any other sequentially sampled periods. Pairwise F_{ST} values were
132 0.0009 and 0.0002 for the 2002-2007 and 2007-2012 collection periods, respectively (Table S3).

133 We used a discriminant analysis of principal components (DAPC; R package adegenet v. 2.1.1,
134 Jombart et al. 2010, Jombart and Ahmed 2011) to further examine temporal changes in the genomic
135 structure of wild *H. zea*. Initially, we identified ddRAD SNP combinations that clustered individuals
136 into groups (k = 2-4) by maximizing between group variation and minimizing within group variation.
137 Using these SNP combinations, we then estimated posterior membership probabilities for each
138 individual from our dataset from each collection year. When k was equal to 2, samples within each of
139 the collection years 2002, 2007, and 2012 were assigned to both genetic clusters. In contrast, samples
140 collected in 2016 were strongly biased toward a single genetic cluster (Figure 1), and this trend

141 persisted regardless of the k-value used (Figure S3). Interestingly, the genotypic shift between 2012
142 and 2016 corresponded to the emergence of widespread field-evolved phenotypic resistance to Cry
143 proteins in *H. zea* (Dively et al. 2016, Kaur et al. 2019). We sequentially removed the top 1%, 2.5%
144 and 5% of ddRAD SNPs that had the greatest influence on genetic clustering, and re-evaluated
145 population membership probability (assuming k=2). From this we determined that up to 5% (n = 720)
146 of ddRAD SNPs were responsible for this observed genotypic shift (Figure S4).
147 To identify which of the 14,398 ddRAD SNPs experienced greater than expected shifts in allele
148 frequency over time, we conducted a genomic outlier analysis (R package OutFLANK v. 0.2, Whitlock
149 and Lotterhos 2015). When we calculated genetic divergence (F_{ST}) for each SNP across all 4
150 collection years, 53 had higher than expected divergence (Figure 2, File S1). Twenty-two of those 53
151 were among the top variance contributing SNPs that strongly influenced genetic clustering in our
152 DAPC analysis. The 53 ddRAD SNP outliers were spread across 47 (of 2,975) unique scaffolds, and
153 their among-year F_{ST} values ranged from 0.068 to 0.278.
154
155 *No significant allelic changes within candidate Bt resistance genes*
156 We found ddRAD-seq markers within or near (<56kb) 10 of the 11 previously identified
157 candidate genes for Bt resistance, but none corresponded to outliers found over the time course of the
158 collections (Table 1). To verify this result we conducted Illumina whole genome sequencing (WGS) of
159 35 *H. zea* males collected from the same location in LA in 2002 (n = 13), 2012 (n = 11), and 2017 (n =
160 11) (Table S4). After genome alignment and filtering, this approach yielded 4,999,128 SNPs, and
161 2,622 fell within the 11 candidate Cry resistance genes. We used a 10kb sliding window-averaged
162 Weir and Cockerham's F_{ST} analysis and observed that F_{ST} values calculated for each gene ranged from
163 -0.018 to 0.056 (Table 1; Figure S5). The empirical threshold for statistically significant genetic
164 divergence, generated according to Rubin et al. (2010), was 6 standard deviations ($ZF_{ST} > 6$) from the
165 mean pairwise F_{ST} value in a window. None of the eleven candidate genes contained windows that

166 reached our threshold for statistical significance of $F_{ST} = 0.079$ ($ZF_{ST} > 6$) per 10kb window. Taken
167 together, these results indicated that the Cry resistance in *H. zea* was not caused by changes in the
168 coding regions of any of these eleven candidate genes or their flanking cis-regulatory sequences.
169
170 *Selection led to significant temporal divergence in novel genomic regions*
171 Using our full WGS SNP dataset, we conducted pairwise analyses of genomic divergence
172 between populations collected from 2002, 2012, and 2017 using 5, 10, 20 and 40 kb sliding window-
173 averaged Weir and Cockerham's F_{ST} along the length of the genome (Figure S6). We report on the 40
174 kb sliding window analysis because it captured the major genomic changes that occurred for all sliding
175 window sizes, but reduced the numbers of windows that appeared highly divergent due to few SNPs in
176 the window. Thirty-six genomic windows on 11 different scaffolds had F_{ST} values that exceeded our
177 empirical threshold for statistical significance of $F_{ST} = 0.047$ per 40 kb when 2002 and 2017
178 populations were compared (Figure 3). For paired comparisons of populations collected in 2002 and
179 2012, and 2012 and 2017 there were 55 and 30 genomic windows on 16 and 10 scaffolds with
180 statistically significant genomic divergence, respectively (Figure 3; File S2).
181 Both genetic drift and selection influence the landscape of genomic divergence between
182 populations. Therefore, we quantified the frequency with which our empirical threshold of $F_{ST} =$
183 0.047, calculated from the pairwise comparison of samples from 2002 and 2017, could be reached in a
184 40kb genomic window due to drift alone (Hudson 2002). Genotypic simulations were conducted for a
185 single population sampled at two timepoints, but in the absence of selection. Allele frequency changes
186 within 40kb windows were examined for multiple population sizes (N = 21,000 - 140,000;
187 Supplemental Methods). The proportion of genotypic simulations (of 20,000) that produced $F_{ST} >$
188 0.047 was the probability that our empirical threshold of $F_{ST} = 0.047$ could be exceeded due to genetic
189 drift. Results from our simulations demonstrated this probability was very low (p = 0.0189 - 0) and
190 negatively correlated with present day population size (Figure S7; Table S5).

191 Because the population sizes used in the simulations are likely underestimates of the effective
192 population size of this migratory pest (Jones et al. 2019), it is highly improbable that our empirical F_{ST}
193 threshold could have been exceeded by the action of drift. As a demonstration of its large effective
194 population size, we collected fifth and sixth instar *H. zea* from ears of Obsession I (n = 8) and Cry-
195 expressing Obsession II (n = 8) sweet corn grown in Central Maryland Research and Education Center
196 (CMREC) demonstration plots. The genomes of these individuals, collected in September of 2017,
197 were sequenced with Illumina short reads, and their genotypes compared to samples collected in
198 Bossier Parish, LA, in that same year. After genome alignment and filtering for quality and linkage
199 disequilibrium, we examined 14,697 SNPs in the dataset, which yielded a genome-wide F_{ST} value of
200 0.0004. The bootstrapped 95% confidence intervals were (-0.0002, 0.0008), and their overlap with
201 zero indicated that allele frequencies in the LA and MD populations did not significantly differ from
202 one another.
203
204 *Identification of novel genes under selection*
205 We examined the genes near regions of greatest temporal genomic divergence to determine
206 whether any were associated with detoxification of plant-produced or human-applied toxicants. Our
207 list of candidates included those from families of genes associated with Cry resistance, but also
208 resistance to synthetic insecticides and xenobiotics. Eleven genes, including 3 peptidases, 4
209 choline/carboxylesterases, 2 cytochrome p450s, 1 ABC transporter, and 1 tetraspanin, were near or
210 within regions of significant temporal genetic divergence. No proteases, proteinases, cadherins,
211 voltage-gated channels, alkaline phosphatases, or glutathione-S-transferases were found within 50 kb
212 of regions with elevated divergence. Evidence from a recent study suggests that increasing adoption of
213 Bt crops led to a reduction in the frequency of synthetic insecticide resistance-conferring alleles, likely
214 due to a reduction in synthetic insecticide applications (Fritz et al. 2018). Therefore, we specifically
215 examined two genes known to be involved in pyrethroid resistance. Neither the voltage-gated sodium

216 channel gene (*vgsc*), nor a cytochrome p450 (*cyp337b3*) which underlie pyrethroid resistance in the
217 close *H. zea* relatives, *Chloridea virescens* (previously *Heliothis virescens*; Taylor et al. 1993) and
218 *Helicoverpa armigera* (Joußen et al. 2012), respectively, were near regions of elevated genomic
219 divergence. All genes found within 50kb of genomic windows with higher than expected divergence
220 can be found in File S3.

221 We reasoned that increasing adoption of Bt crops over time would result in strong, sustained
222 selection over time, leading to significant allele frequency changes across multiple collection years.
223 Therefore, we identified scaffolds with elevated temporal genomic divergence in more than one
224 pairwise comparison. Five scaffolds showed signs of significant allele frequency change in at least two
225 pairwise comparisons. These included scaffold 173, which contained a *venom dipeptidyl peptidase 4* -
226 like gene (HzOG204148), and scaffold 569, which contained two *carboxypeptidase Q (cpq)* -like genes
227 (HzOG204151 and HzOG204153). Scaffold 1964 contained an RNA-directed DNA polymerase
228 (HzOG206500), and scaffold 65 had one genomic window of elevated divergence in multiple pairwise
229 comparisons which was near 7 genes, including *tetraspanin* (*tspan*) *68C* (HzOG211510) and a
230 *krueppel*-like gene (HzOG211515). Finally, scaffold 1612 contained 2 genomic windows with
231 elevated divergence across all three by-year comparisons, which overlapped with *cyp333b3*
232 (HzOG200024). Of these five scaffolds, 569 and 1612, had levels of genetic divergence that greatly
233 exceeded all others between 2002 and 2017 ($F_{ST} > 0.089$), and such divergence was never observed for
234 genetic simulations in the absence of selection (Table S5).

235 The genome assembly of *H. zea* is fragmented into 2,975 genomic scaffolds, which makes
236 detecting the extent of a selective sweep challenging. Under conditions of strong selection, the
237 footprint in the genome should be broad, and we reasoned there was potential for some of these
238 temporally divergent scaffolds to be physically linked. We used tblastn to align the genes from the 5
239 scaffolds showing strong temporal divergence in multiple by-year comparisons to the *Bombyx mori*
240 genome (ASM15162v1), and identified their relative positions assuming conserved macrosynteny

among Lepidoptera (d'Alençon et al. 2010). Four of these five scaffolds were adjacent to one another and syntenic to *B. mori* chromosome 13, indicating that the breadth of the selective sweep extended for *ca.* 3 Mb (Fig S8). Linkage disequilibrium within the putative sweep region decayed more slowly than both genome-wide average rates of decay, as well as decay along single, similarly-sized scaffolds (> 1Mb) without evidence of genomic divergence (Fig S9). This region of the *B. mori* genome contained 134 unique genes, none of which were previously known to be involved in Cry resistance (File S4). Interestingly, we re-examined our ddRAD-seq dataset and detected two ddRAD-seq markers found within this 3 Mb genomic region (on scaffold 569), which were outliers according to OutFLANK. This further confirmed the significance of changes occurring in this genomic region, but with a much larger sample size (n > 45 per year).

The 3 Mb putative sweep region experienced the greatest temporal change in frequency according to our whole genome scan. Yet our DAPC analysis identified 720 ddRAD SNPs on 505 unique scaffolds which caused a major architectural shift in the genomes of *H. zea* collected between 2012 and 2016 (Figure 1, Figure S3, Figure S4). Of those 720 ddRAD SNPs, only 5 aligned to this 3 Mb region. This pattern, consistent with polygenic selection on standing variation, indicated that allele frequency changes at many loci are involved in *H. zea* adaptation over the study period. Our further analyses in this study focused on the 3 Mb putative sweep region due to the strong changes in allele frequency and the breadth of the putative selective sweep.

Temporal shifts in the frequency of non-synonymous mutations

Given the significant levels of genomic divergence near *cyp333b3* and the two *carboxypeptidase Q-like (*hereafter *cpq*) genes detected by ddRAD-seq and WGS datasets, as well as the putative functional roles of these genes in insect metabolism and detoxification, we further examined protein-coding changes at these genes over time. Based upon our temporal analysis, genomic divergence was strongest at the 3' end of all three genes (Figure S10). While we detected 50

266 and 40 SNPs within *cpq1* and *cpq2*, respectively, only 3 were associated with putative changes in
267 amino acid sequence (Table 3). One transition from cytosine to thymine caused a putative alanine to
268 valine substitution in *cpq1*, and over time, the cytosine declined significantly from 0.73 in 2002, to 0.32
269 and 0.14 in 2012 and 2017, respectively (Freeman-Halton Fisher's exact $p = 0.0001$; Table 2). Two
270 SNPs in *cpq2* showed modest changes in allele frequency over time, although not statistically
271 significant when a Bonferroni-corrected α value was used to account for multiple comparisons (Table
272 2). One cytosine to thymine transition encoded an alanine to valine amino acid substitution at 152,595
273 bp, and a second occurred at 152,183 bp, encoding a threonine to isoleucine substitution. Finally, there
274 were 28 SNPs present in the protein-coding sequence of *cyp333b3*, 10 of which produced putative
275 amino acid changes. Five of these non-synonymous mutations underwent statistically significant
276 changes over time ($p < 0.003$; Table 2).
277 We also compared the frequencies of these non-synonymous mutations in the LA and MD
278 populations collected in 2017, and for all 8 SNPs, there was no statistically significant difference in
279 allele frequency based on collection location (Table S7). These results and the low overall F_{ST} value
280 for the two-population comparison (reported above) support previous findings that *H. zea* is highly
281 migratory with little population structure.
282
283 *Strength of selection experienced by wild* H. zea
284 Using the temporal changes in allele frequency, we calculated the strength of selection required
285 to produce these statistically significant protein-coding changes over time. The selection coefficient(s)
286 gave the strength of selection against q, the declining allele, examined under 3 different dominance
287 scenarios of p: complete dominance, no dominance, and recessiveness. For 5 of the 8 non-synonymous
288 mutations (Table 2), q was the allele present in the reference genome, and in the remaining cases, q was
289 the alternate allele. We calculated mean values of s from allele frequency changes which occurred
290 during the 2002-2012 collection period, as well as the 2012-2017 collection period. Values of s varied

291 according to selection period and dominance scenario (Figure 4), but most fell within the range of 0.03-
292 0.17. When *p* was assumed to be recessive, stronger selection was required to produce the observed
293 changes in allele frequency in the 2002-2012 collection period due to its low frequency (*ca.* 0.25 - 0.3).
294 Yet when *p* was dominant, stronger selection was required to produce the observed changes in allele
295 frequency between 2012-2017, due to its relatively high starting frequency (*ca.* 0.65-0.75). In the case
296 of the SNP at 36,959 bp on scaffold 1612, it's frequency (*p*) was already 0.95 in 2012. Assuming
297 dominance of *p*, the strength of selection that would have been required to increase its frequency to 1
298 by 2017 was very high ($s = 0.70$), which increased the variance around the mean value of *s* (Figure 4).
299

300 *Frequency of non-synonymous substitutions prior to 2002*

301 The intermediate frequency of the non-synonymous polymorphisms in 2002 compelled us to
302 examine the frequencies of those polymorphisms in *H. zea* collected from our LA trapping sites in
303 1998, shortly after the release of Cry-expressing crops in 1996. Moths from 1998 had DNA quality
304 that was too low for use in ddRAD-seq and WGS analyses, but short fragments could be PCR
305 amplified. We therefore amplified and Sanger sequenced *ca.* 200bp DNA fragments spanning one
306 non-synonymous polymorphism in each of *cpq1* and *cyp333b3* from 20-26 *H. zea* collected in 1998.
307 The respective frequencies of the reference alleles for *cpq1* and *cyp333b3* were 0.93 and 0.90 in 1998
308 compared with 0.73 and 0.65 in 2002, and 0.14 and 0.18 in 2017 (Table 2) showing further evidence of
309 genomic divergence at these sites. The change in allele frequency between 1998 and 2002 occurred
310 prior to introduction of varieties with Cry2 genes.
311

312 *Strongest selective sweep is not linked to Cry resistance*

313 To test whether this 3 Mb region with significant allele frequency divergence was linked to Cry
314 resistance, we bioassayed the progeny of two F_2 genetic crosses between recently field-collected Cry

315 resistant grandparents and Cry susceptible grandparents (Benzon Research Inc., Carlisle, PA). Cry
316 resistant larval growth was 7 and 10 times greater than that of Benzon susceptibles when fed a
317 diagnostic dose of diet containing Cry1 and Cry1 + Cry2 sweetcorn leaf tissue, made according to
318 Dively et al. (2016). Following 7 days of exposure to Cry1 or Cry1 + Cry2 treated diet, F_2 larval
319 growth phenotypes from each family were measured and individuals genotyped by PCR at non-
320 synonymous mutations in the *cyp333b3* and *cpq1* genes. A Kruskal Wallis test found no significant
321 difference in larval growth according to genotype on either diet treatment, indicating these non-
322 synonymous mutations were not additively associated with resistance to Cry toxins (Figure 5; Table
323 S9). For the *H. zea* system, the question of what environmental factor could have selected for such
324 rapid change in allele frequency is fascinating but unsolved.

325

326 **Discussion**

327 In this study, we compared the genomes of *H. zea* collected between 2002 and 2017 to quantify
328 changes that occurred during a period of rapid adoption of a major agricultural innovation, Cry toxin-
329 expressing crops, in the United States. We found substantial changes in the frequencies of SNPs
330 throughout the *H. zea* genome that occurred in concert with the evolution of resistance to Cry toxins in
331 field populations of *H. zea*. A major finding of our work is that none of the previously identified
332 candidate Cry resistance genes in Lepidoptera showed signatures of selection in the field-collected *H.*
333 *zea*, nor were genes in the strongest selective sweep region associated with resistance in an additive
334 way. Instead, the timing and complexity of the genomic changes that occurred in *H. zea* were
335 consistent with the predicted polygenic nature of Cry resistance in this species, and provide insight into
336 how, when and whether genomic approaches could be used for resistance risk monitoring.

337 Over the past four decades there has been an emphasis on the search for single genes that confer
338 resistance to control measures in insect pests and weeds, and we now have a catalogue of such genes

339 (Ffrench-Constant 2013, Feyereisen et al. 2015, Délye et al. 2013, Heckel 2020). There are cases
340 where a single gene diagnostic (Donnelly et al. 2016) or a genomic diagnostic targeted at multiple
341 previously established candidate genes (Weetman et al. 2018) has been useful in monitoring for
342 resistance, but recently there has been concern that these focused approaches could lead to
343 complacency, as resistance evolves based on unexpected genes (Heckel 2020) or based on multiple
344 genes (Délye et al. 2013, Van Leeuwen et al. 2020, Leon et al. 2020) that are already at low
345 frequencies in pest or weed populations. Our work demonstrates that these concerns are appropriate. If
346 any or all of the previously identified candidate genes had been used in a diagnostic for Cry resistance
347 in *H. zea*, the diagnostic would have missed the coming problem.

348 While functional roles of specific genes that we identified have not yet been demonstrated,
349 multiple *H. zea* genomic regions have changed dramatically since the introduction of Cry toxin-
350 producing cotton and corn. The most compelling changes were found within a broad selective sweep
351 on the *H. zea* homolog of *B. mori* chromosome 13. WGS and ddRAD-seq datasets both showed
352 stronger than expected temporal genomic changes in a roughly 3 Mb region over the 15 year sampling
353 period. When the results of the genome analyses are combined with PCR analysis of SNPs within
354 candidate genes from 1998 samples from the same location, allele frequency changes of up to 0.79
355 occurred in these genes with metabolic functions or that are associated with xenobiotic resistance
356 (Ward 1976, Guzov et al. 1998, Bown and Gatehouse 2004). Analysis of genetic crosses of resistant
357 and susceptible strains found that this genomic region was not linked to Cry resistance in an additive
358 way (Figure 5). This was surprising due to rapid allele frequency changes, and the fact that two of the
359 candidates (*cpq1* and *cpq2*) were genes from a family capable of binding Cry toxins in *H. armigera*
360 (Da Silva et al. 2018). While this does not eliminate the possibility that genes in this region have
361 epistatic interactions that confer resistance, it highlights that monitoring solely for strong selective

362 sweeps could potentially mislead when it comes to resistance gene identification and resistance risk
363 monitoring.

364 We emphasize, however, that we were able to detect that multiple regions of the genome
365 changed during the course of our sampling in a pattern consistent with polygenic adaptation from
366 standing genetic variation. Our findings are consistent with a recent study of crosses of Cry2 resistant
367 *H. zea* to a susceptible strain that found phenotypic variance fitting a polygenic model (Yang et al.
368 2020). In our study, strong polygenic change was most apparent between the years 2012 and 2016
369 (Figure 1) when widespread and damaging levels of resistance to Cry toxins were documented (Dively
370 et al. 2016). Ongoing, experimental, quantitative genomic studies will shed light on which of these
371 genomic regions contribute most substantially to Cry tolerance in field-collected larvae.

372 The use of genomic approaches to study rapid evolution in natural populations has gained
373 momentum (*e.g.* Crossley et al. 2017, Fritz et al. 2018, Nosil et al. 2018, Weetman et al. 2018, Bi et al.
374 2019, Dayan et al. 2019, Laurentino et al. 2020). Yet their use for studying pest adaptation to
375 agricultural systems has been limited to a few cases (reviewed in Pélissié et al. 2018). As we attempt to
376 apply these approaches to a variety of agricultural systems, system- and method-specific issues will
377 need to be addressed to ensure accuracy. Results from our experiments indicate that genomic
378 approaches may hold promise for monitoring resistance risk in pest populations exposed to novel
379 management practices, with several caveats for how we interpret our results.

380 One concern is that major genomic changes identified through genomic monitoring might not
381 be due to the control tool being studied. Deployment of pest management innovations do not take
382 place in the absence of other existing selection pressures in the environment. When transgenic cotton
383 and corn were introduced into North American agricultural systems, the use of synthetic insecticides
384 were reduced, though certainly not eliminated. In a previous study using ddRAD-seq, we found that

385 another pest of cotton in Louisiana, *Chloridea virescens*, did not experience changes in the frequency
386 of Bt resistance candidate genes. Rather, changes to the frequency of an allele that conferred resistance
387 to pyrethroid insecticides were observed over a 15 year period (Fritz et al. 2018). Indeed, our data
388 suggest that the greatest allele frequency changes in the *H. zea* genome were not directly associated
389 with Cry resistance. Therefore, confirming the strength of association between any major genomic
390 change with a resistance phenotype is critical.

391 This confirmation testing can be conducted using the types of simple crosses and PCR-based
392 assays that we carried out in our study. Such assays can falsify the importance of rapidly changing loci
393 to resistance phenotypes and alleviate concerns regarding resistance risk. Conversely, knowledge of
394 allele frequency changes and confirmation of their impact on field resistance could then be used to
395 make predictions for future changes in resistance in field populations of pests. In cases similar to *H.*
396 *zea* on Cry-expressing crops where the control tactic does not kill all pest individuals or has sublethal
397 effects, collections from paired sentinel plots could also point to genomic regions associated with
398 resistance. Such paired sentinel plots are already being used for determining the current extent of
399 resistance, but typically monitor damage to the crop rather than the genetics of the pest (Dively et al.
400 2016). Focusing on damage and the economic impact of resistance does little to predict risk of
401 resistance and then mitigate it. Once widespread damage has occurred, resistance alleles have likely
402 already spread widely in the landscape. Instead, sampling and sequencing early instar larvae found in
403 treated and untreated sentinel plots would allow researchers to respectively quantify genomic change
404 associated with the pest management tool compared to the population at large.

405 A second concern is that population demographic processes, including migration patterns and
406 population expansion can result in the appearance of genome-wide shifts in allele frequency, which are
407 not due to selection. When this occurs, the risk for genomic monitoring would be to falsely implicate a
408 region of the genome in evolving resistance. We were fortunate to be studying a highly migratory

409 species (Hartstack et al. 1982, Gould et al. 2002, Seymour et al. 2016) with little population structure
410 over large geographic areas (Perera and Blanco 2011, Seymour et al. 2016). Our finding of negligible
411 genomic differences between samples collected in 2017 from Louisiana and Maryland are consistent
412 with results from these previous studies. Additionally, *H. zea* feeds on a number of host plants, and this
413 seems to have resulted in the population densities remaining similar to those prior to the use of
414 transgenic cotton and corn (Micinski et al. 2008). It is not surprising that with this species, measured
415 overall genomic diversity remained constant. For a species that interbreeds over such a large area,
416 such population demographic processes are unlikely to be an issue. Had we been studying an insect or
417 weed with strong local population structure, or one in which the control measure caused severe
418 reduction and resurgence of the pest, our experimental approaches would have required additional
419 careful consideration.

420 Finally, many so-called "mega-pest" lineages are thought to have genomes that already contain
421 variants, which prime them for the evolution of resistance to novel pest control measures (Crossley et
422 al. 2017, Pélissié et al. 2018). When selection is applied to these populations, they experience soft
423 selective sweeps at genomic regions that are typically spread throughout the genome. Such signatures
424 of polygenic selection in field-collected samples can be difficult to bioinformatically tease apart from
425 the impacts of drift. Fortunately, new computational approaches which can better differentiate soft
426 selective sweeps from background selection or other population demographic processes are becoming
427 available (Schrider and Kern 2016), and some even detect the proportion of genome-wide changes in
428 populations due to selection using temporal datasets (Buffalo and Coop 2019). Further development of
429 bioinformatic approaches specifically aimed at detecting temporal genomic changes at both single loci
430 and those spread throughout the genome will provide critical tools for resistance risk monitoring in
431 agricultural systems.

The benefits of genomic monitoring of pest species go beyond identifying the mechanisms of resistance and applications for resistance risk mitigation. Adoption of genomic approaches for resistance monitoring will be crucial for addressing the decades-long debate on exactly how pest species rapidly adapt to management practices (Crow 1957). Even today, the extent to which a species response to selection is mono- or polygenic, from standing genetic variation or novel mutation remains unclear. Here we present evidence on the nature of genomic change during expansion of a major agricultural innovation, Cry-expressing crops, in one important pest species. Extension of genomic monitoring programs, accompanied by targeted experimental confirmation, to additional pest species on broad spatial and temporal scales could provide critical insight into the nature of adaptation by these species to human-imposed selection.

Data Availability

Data for this study are available at: to be completed after manuscript is accepted for publication.

All scripts and barcodes files used to perform these analyses can be found at https://github.com/mcadamme/FieldHz_Pop_Genomics.

Competing Interests

The authors declare they have no competing interests.

Acknowledgements

This work was funded by USDA NIFA Biotechnology Risk Assessment Grants 2016-33522-25640 and 2019-33522-29992.

457 **Methods**

458 *Insect collections for ddRAD-seq and WGS analyses*

459 *H. zea* adults were collected by pheromone-baited trap bi-weekly from May through September
460 in Bossier Parish, Louisiana (Table S1). Collections were made in the years 2002, 2007, 2012, 2016,
461 and 2017, and specimens were stored in 100% ethanol at -20° C until use. Late instar (5-6th) *H. zea*
462 were also collected in August and September of 2017 at the University of Maryland CMREC farms in
463 Prince George's county, MD, from the ears of Bt-expressing and non-Bt sweet corn isolines according
464 to Dively *et al.* (2016) and reared to pupation under laboratory conditions of 16:8 LD at 25° C and 50%
465 RH. Newly emerged adults were sacrificed by freezing and held at -80° C until use. Genomic DNA
466 was isolated from half of an adult thorax with a Qiagen DNeasy kit (Qiagen, Inc., Valencia, CA, USA)
467 using a modified mouse tail protocol according to Fritz et al. (2020).

468

469 *DdRAD-seq library preparation*

470 Specimens from the years 2002, 2007, 2012, and 2016 were prepared into double-digest RAD-
471 seq (ddRAD-seq) libraries for population genomic analysis according to Fritz *et al*. (2016 2018; further
472 details in SI). Barcoded and indexed libraries were combined in equimolar amounts (4 nM) for
473 sequencing on 8 runs of an Illumina MiSeq. Our final data set contained 259 individuals after quality
474 control and filtering to remove individuals with low read counts (Figure S1).

475

476 *DdRAD-seq Bioinformatic Analysis*

477 After adapter trimming (cutadapt v. 1.18; Martin, 2011), paired end reads were merged with
478 FLASh (v. 1.2.11; Magoc and Salzberg, 2011), demultiplexed and filtered using the process_radtags
479 script from Stacks (v. 2.2; Catchen et al. 2011), and aligned to the *H. zea* genome assembly (v. 1.0,
480 NCBI Bioproject PRJNA378438; Pearce et al., 2017) with Bowtie2 (v. 2.2.6-2; Langmead and
481 Salzberg, 2012). We called genotypes with BCFtools mpileup and call functions (Danecek et al., 2016)

482 using standard options. Our final SNP dataset was filtered with vcftools (Danecek et al., 2011) to
483 include SNPs that i) had a depth of coverage > 3, ii) were present in at least 75% of the 259 individuals
484 in our filtered dataset, iii) had a minor allele frequency of 0.05, iv) had a maximum of 2 alleles, and v)
485 were thinned to 1 per 200 bp to reduce the degree of linkage disequilibrium between SNPs on the same
486 ddRAD-seq marker.

487 We calculated average nucleotide diversity (π) and heterozygosity (the inbreeding coefficient
488 F) values and their corresponding 95% bootstrapped confidence intervals (N = 5000) for each year
489 using vcftools. Genome wide allele frequency divergence between years was examined by pairwise
490 F_{ST} (Weir and Cockerham 1984) using the StAMPP package (v. 1.5.1, Pembleton *et al.,* 2013) in R (v.
491 3.4.4, R Development Core Team, 2008). A k-means analysis, implemented with adegenet (v. 2.1.1,
492 Jombart and Ahmed, 2011), modeled the number of putative populations or clusters (k = 1:10) into
493 which our *H. zea* fit according to genotype. Using the three best-fitting numbers of clusters (k = 2:4)
494 as priors, we applied a Discriminant Analysis of Principal Components (DAPC; adegenet) to identify
495 groups of genetically similar *H. zea* individuals. The R package OutFLANK (v. 0.2, Whitlock and
496 Lotterhos 2015) calculated Weir and Cockerham's F_{ST} for each SNP to quantify genetic divergence
497 across years and identify those with higher than expected divergence over time.

498 A number of genes have already been implicated in Bt resistance in other Lepidopteran pest
499 species (Table 1). Blastn and tblastn searches (https://blast.ncbi.nlm.nih.gov) determined the locations
500 of these putative Bt resistance genes in the *H. zea* genome, and we used custom scripts to identify the
501 nearest ddRAD-seq marker to each (Table 1).

502

503 *WGS library preparation*

504 To confirm our ddRAD-seq results, and extend our analysis to regions not covered by ddRAD-
505 seq markers, we conducted whole genome sequencing of additional *H. zea* collected from this same
506 location in LA. Genomic DNA from additional adults collected in 2002 (n= 13), 2012 (n = 11), and

507 2017 (n = 11) was prepared according to an Illumina TruSeq LT library preparation (Illumina, Inc. San
508 Diego, CA), and indexed libraries were sequenced on two Illumina NextSeq500 runs using 150 bp
509 paired-end reads.
510
511 *WGS bioinformatic analysis*
512 Following filter-trimming (Trimmomatic v. 0.38; Bolger et al., 2014), and reference genome
513 alignment with Bowtie2, we filtered poor quality and duplicate (PCR and optical) reads from our
514 dataset using GATK (v. 4.0.12.0; McKenna et al. 2010) and SAMtools (v. 0.1.19; Li et al., 2009).
515 Genotypes were called with BCFtools mpileup and call functions, and indels were removed with
516 vcftools. Further filtering with vcftools removed SNPs i) with a coverage depth < 3, ii) absent from 8
517 or more individuals sequenced, iii) with minor allele frequency of < 0.05, and iv) with > 2 alleles per
518 locus. Genomic change at the 11 Bt resistance candidate genes was measured by pairwise F_{ST}
519 averaged across SNPs found within each gene for samples collected in 2002 and 2017. According to
520 Rubin et al. (2010), we standardized F_{ST} estimates with a z-transformation, and excluded windows with
521 fewer than 10 SNPs. Windows with an average F_{ST} value greater than 6 standard deviations from the
522 genome-wide averaged F_{ST} value ($ZF_{ST} > 6$) were considered to have undergone statistically significant
523 genetic divergence. For each scaffold containing a Bt resistance candidate, 10kb averaged F_{ST} values
524 with a 1kb step size were examined for the length of the scaffold and compared against our empirical
525 threshold determined for each genome-wide by-year comparison.
526 Sliding window averaged F_{ST} values were used to identify signatures of strong selection across
527 the entire genome. For each pair of years we calculated averaged F_{ST} across 40kb windows with a
528 10kb step size. Additional sliding window and step sizes were also examined for the comparison of
529 samples collected in 2002 and 2017, and are reported in Figure S2. Genotype simulations under
530 neutral expectations using realistic population demographic conditions generated a distribution of F_{ST}
531 values, which were compared against empirical F_{ST} values from our genome scan results as a

532 secondary approach to examining statistical significance (Hudson 2002; Supplemental Methods).
533 Results of our whole genome scan were compared to results from our ddRAD-enabled outlier analysis
534 to determine which, if any, regions were detected in both datasets.

535 We examined the genes within 50kb of each genomic window showing signs of statistically
536 significant genetic divergence between each pair of years, and queried the gene list to identify those
537 from families generally known to be associated with detoxification of plant-produced or human-applied
538 toxicants. These included peptidases/proteases (Gill et al. 1995, Banks et al. 2001, Liu et al. 2015),
539 carboxyl/cholinesterases (Newcomb et al. 1997, Russell et al. 2011), proteinases (Ahmed et al. 1998),
540 cytochrome p450s (reviewed in Russell et al. 2011, Howe and Herde 2015), cadherins (Gahan et al.
541 2001, Morin et al. 2003, Fabrick et al. 2014; Zhang et al. 2017), ABC transporters (Gahan et al. 2010),
542 alkaline phosphatases (Jurat-Fuentes and Adang 2004), glutathione-S-transferases (reviewed in Russell
543 et al. 2011, Howe and Herde 2015), voltage-gated sodium channels (reviewed in Dong et al. 2014), and
544 tetraspanins (Jin et al. 2018).

545 The genome assembly of *H. zea* is fragmented into 2,975 genomic scaffolds, which makes
546 detecting the extent of a selective sweep challenging. A tblastn alignment of genes from the 5 most
547 divergent scaffolds to the *Bombyx mori* genome (NCBI ASM15162v1) identified their relative
548 positions to one another using conserved macrosynteny among Lepidoptera (d'Alençon et al. 2010).
549

550 *Genome similarity between* H. zea *collected from MD and LA in 2017*

551 A global F_{ST} analysis was implemented using the R package hierfstat (v. 0.04-22; Goudet
552 2005), and compared *H. zea* collected at CMREC in Prince George's county, MD, in 2017 to those
553 collected in Bossier Parish, LA, in 2017. Bootstrapped 95% confidence intervals around the global F_{ST}
554 value were generated using the R package bootstrap (v. 2019.6; Tibshirani and Efron 1993). Overlap
555 of the 95% confidence intervals with zero indicated a lack of statistical significance.
556

557 *Gene-specific analyses*

558 Putative protein-coding changes were examined for three genes found within regions of highest
559 genetic divergence across years, *cpq1*, *cpq2* and *cyp333b3*. We extracted the nucleotide sequence for
560 the coding region of each gene using a custom bioinformatic pipeline, and ExPASy
561 (https://web.expasy.org/translate/) was used to produce the putative protein sequences from individuals
562 collected in 2002 and 2017. Protein sequences were aligned to one another using MUSCLE (Edgar
563 2004) to identify putative amino acid substitutions. Allele frequencies for SNPs which caused these
564 substitutions were calculated for each collection year. A two-sided 3x2 Freeman-Halton extension of
565 the Fisher's exact tests with a Bonferroni-corrected alpha value ($\alpha = 0.006$) was used to examine
566 whether allele frequencies differed significantly among years. Frequencies of these 8 non-synonymous
567 SNPs were also compared in 2017 collections of *H. zea* from MD and LA using this same approach to
568 quantify the extent of similarity over geographical distance.

569

570 *Strength of selection*

571 The strength of selection (s) imposed upon ancestral codons in the protein-coding regions of
572 *cpq1*, *cpq2*, and *cyp333b3* were examined over time (Falconer and Mackay 1996). We calculated s for
573 two periods of selection: 2002 - 2012, as well as 2012 - 2017. We assumed 5 generations per year
574 (Reay-Jones 2019) for each selection period, and calculations were performed assuming dominance of
575 the allele that was increasing in frequency, p, no dominance of p, and recessiveness of p.

576

577 *Frequency of non-synonymous substitutions prior to 2002*

578 Genomic DNA was isolated from *H. zea* collected from our LA traps in 1998 using a Qiagen
579 DNeasy isolation kit as indicated above. DNA from each individual was PCR amplified using primers
580 flanking each non-synonymous mutation (Table S8), which allowed us to determine the frequency of
581 these mutations prior to 2002. Amplicons were sequenced on an ABI3730xl (Applied Biosystems,

582 Foster City, CA, USA), and chromatograms were visually inspected for quality filtering and sample
583 genotyping.
584
585 *Test for linkage*
586 In 2019, we crossed a population of Cry resistant *H. zea* initiated from late instars collected out
587 of Cry1 + Cry2 expressing sweet corn ears to Cry susceptible *H. zea* (Benzon Research Inc., Carlisle,
588 PA). Following 7 days of exposure to Cry1 or Cry1 + Cry2 treated diet (Supplemental Methods),
589 larval growth phenotypes for F2s from each family were scored and individuals genotyped at non-
590 synonymous mutations in the *cyp333b3* and *cpq1* genes. For genotyping, DNA from each individual
591 was PCR amplified using primers in Table S8 and genotyped using a Sanger approach, as above.
592 Associations between larval genotype and Cry resistance phenotype were analyzed using ANOVA.

593 **References**

594 Ahmed S, Wilkins RM, Mantle D (1998) Comparison of proteolytic enzyme activities in adults of
595 insecticide resistant and susceptible strains of the housefly *M. domestica* L. *Insect Biochem Mol*
596 *Biol* 28(9): 629-639.

597 Bagwell RD, Graves JB, Leonard BR, Burris E, Micinski S, Fife JH, Mascarenhas V (1996) Status of
598 insecticide resistance in tobacco budworm and bollworm in Louisiana during 1995. In Proceedings
599 Beltwide Cotton Prod. Conf., National Cotton Council, Memphis, TN. pp. 1059-1067

600 Banks DJ, Jurat-Fuentes JL, Dean DH, Adang MJ (2001) *Bacillus thuringiensis* Cry1Ac and Cry1Fa
601 delta-endotoxin binding to a novel 110kDa aminopeptidase in *Heliothis virescens* is not N-
602 acetylgalactosamine mediated. *Insect Biochem Mol Biol* 31: 909-918.

603 Baucom RS (2019) Evolutionary and ecological insights from herbicide-resistant weeds: what have we
604 learned about plant adaptation, and what is left to uncover? *New Phytol* 223:68-82.

605 Bi K, Linderoth T, Singhal S, Vanderpool D, Patton JL, Nielsen R, et al. (2019) Temporal genomic
606 contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change.
607 *PLoS Genet* 15(5): e1008119. https://doi.org/10.1371/journal.pgen.1008119

608 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data.
609 *Bioinformatics*, btu170.

610 Bown DP and Gatehouse JA (2004) Characterization of a digestive carboxypeptidase from the insect
611 pest corn earworm (*Helicoverpa armigera*) with novel specificity towards C-terminal glutamate
612 residues. *EJB* 271(10): 2000-2011.

613 Catchen J, Amores A, Hohenlohe P, Cresko W, Postlethwait J (2011). *Stacks: building and genotyping*
614 *loci de novo from short-read sequences*. *G3* 1:171-182.

615 Crossley MS, Chen YH, Groves RL, Schoville SD. Landscape genomics of Colorado potato beetle
616 provides evidence of polygenic adaptation to insecticides (2017). Mol Ecol 26(22):6284-6300. doi:
617 10.1111/mec.14339.

618 Crow JF (1957) Genetics of insect resistance to chemicals. *Ann Rev Entomol* 2: 227-246.

619 d'Alençon E, Sezutsu G, Legeai F, Permal E, Bernard-Samain S, Gimenez S, Gagneur C, Cousserans
620 F, Shimomura M, Brun-Barale A, Flutre T, Couloux A, East P, Gordon K, Mita K, Quesneville H,
621 Fournier P, Feyereisen R (2010) Extensive synteny conservation of holocentric chromosomes in
622 Lepidoptera despite high rates of local genome rearrangements. *PNAS* 107(17): 7680-7685; DOI:
623 10.1073/pnas.0910413107

624 Da Silva IHS, Goméz I, Sánchez J, Martínez de Castro DL, Valicente FH, Soberón M, et al. (2018)
625 Identification of midgut membrane proteins from different instars of *Helicoverpa armigera*
626 (Lepidoptera: Noctuidae) that bind to Cry1Ac toxin. *PLoS ONE* 13(12): e0207789.
627 https://doi.org/10.1371/journal.pone.0207789

628 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker R, LUnter G, Marth
629 G, Sherry ST, McVean G, Durbin R and 1000 Genomes Project Analysis Group (2011) The
630 Variant Call Format and VCFtools. *Bioinformatics* 27(15): 2156–2158.

631 Danecek P, Schiffels S, Durbin R (2016) Multiallelic calling model in bcftools (-m),
632 samtools.github.io/bcftools/call-m.pdf

633 Délye C, Jasieniuk M, Le Corre V (2013) Deciphering the evolution of herbicide resistance in weeds.
634 *Trends Genet* 29:649-658.

635 Dively GP, Venugopal PD, Finkenbinder C (2016) Field-Evolved Resistance in Corn Earworm to Cry
636 Proteins Expressed by Transgenic Sweet Corn. *PLoS ONE* 11(12): e0169115.

637 Dively G.P., P. D. Venugopal, D. Bean, J. Whalen, K. Holmstrom, T.P. Kuhar. H.B. Doughty, T.
638 Patton, B. Cissel, and W.D. Hutchison. (2018) Regional pest suppression associated with
639 widespread Bt maize adoption benefits vegetable growers. *Proc Natl Acad Sci* 115:3320-3325.

640 Dong K, Du Y, Rinkevich F, Nomura Y, Xu P, Wang L, Silver K, and Zhorov BS (2014) Molecular
641 biology of insect sodium channel and pyrethroid resistance. *Insect Biochem Mol Biol* 50: 1-17.

642 Donnelly MJ, Isaacs A, Weetman D (2016) Identification, validation, and application of molecular
643 diagnostics for insecticide resistance in malaria vectors. *Trends Parasitol* 32:197-206.

644 Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput.
645 *Nuc Acids Res 32*(5):1792–1797. https://doi.org/10.1093/nar/gkh340

646 Fabrick JA, Ponnuraj J, Singh A, Tanwar RK, Unnithan GC, Yelich AJ, Li X, Carrière Y, Tabashnik
647 BE (2014) Alternative splicing and highly variable cadherin transcripts associated with field-
648 evolved resistance of pink bollworm to Bt cotton in India. *PLoS One* 9(5): e97900.

649 Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics. 4th Edition, Prentice Hall,
650 Essex.

651 Feyereisen R, Dermauw W, Van Leeuwen T (2015)Genotype to phenotype, the molecular and
652 physiological dimensions of resistance in arthropods. *Pestic Biochem Physiol* 121:61-77.

653 Ffrench-Constant RH (2013) The molecular genetics of insecticide resistance. *Genet* 194(4):807-815.

654 Fritz ML, Paa S, Baltzegar J, Gould F (2016) Application of a dense genetic map for assessment of
655 genomic responses to selection and inbreeding in *Heliothis virescens*. *Insect Mol Biol* 25(4): 385-
656 400.

657 Fritz ML, DeYonke AM, Papanicolaou A, Micinski S, Westbrook J, Gould F (2018) Contemporary
658 evolution of a Lepidopteran species, *Heliothis virescens*, in response to modern agricultural
659 practices. *Mol Ecol* 27(1): 167-181.

660 Fritz ML, Nunziata SO, Guo R, Tabashnik BE, Carrière Y (2020) Mutations in a novel cadherin
661 gene associated with Bt resistance in *Helicoverpa zea*. *G3* 10(5):1563-1574.

662 Gahan LJ, Gould F, Heckel DG (2001) Identification of a gene associated with Bt resistance in
663 *Heliothis virescens*. *Sci* 293(5531): 857-60.

664 Gahan LJ, Gould F, Lopez JD, Micinski S, and Heckel DG (2007) A polymerase chain reaction screen
665 of field populations of *Heliothis virescens* for a retrotransposon insertion conferring resistance to
666 *Bacillus thuringiensis* toxin. *J Econ Entomol* 100: 187-194.

667 Gahan LJ, Pauchet Y, Vogel H, Heckel DG (2010) An ABC transporter mutation is correlated with
668 insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLoS Genet* 6: e1001248.

669 Gill S, Cowles EA, Francis V (1995) Identification, isolation, and cloning of a *Bacillus thuringiensis*
670 Cry1Ac toxin-binding protein form the midgut of the lepidopteran insect, *Heliothis virescens*. *J*
671 *Biol Chem* 270: 27277-27282.

672 Goudet J (2005) hierfstat, a package for r to compute and test hierarchical *F*-statistics. *Mol Ecol Notes*
673 5(1):184-186.

674 Gould F (1991) The evolutionary potential of crop pests. *Am Sc* 79: 496-507.

675 Gould F (1998) Sustainability of transgenic insecticidal cultivars: integrating pest genetics and ecology.
676 *Annu Rev Entomol* 43: 701–726.

677 Gould F, Blair N, Reid M, *et al*. (2002) *Bacillus thuringiensis*-toxin resistance management: stable
678 isotope assessment of alternate host use by *Helicoverpa zea*. *Proc Natl Acad Sci USA* 99:16581–
679 16586.

680 Gould F, Brown ZS and Kuzma J (2018) Wicked evolution: Can we address the sociobiological
681 dilemma of pesticide resistance? *Sci* 360(6390): 728-32.

682 Guo Z, Kang S, Zhu Z, Zia J, Wu Q, Wang S, Xie W, Zhang Y (2015a) Down-regulation of a novel
683 ABC transporter gene (Pxwhite) is associated with Cry1Ac resistance in the diamondback moth,
684 *Plutella xylostella* (L.) *Insect Biochem Mol Biol* 59: 30-40.

685 Guo Z, Kang S, Chen D, Wu Q, Wang S, Xie W, et al. (2015b) MAPK Signaling Pathway Alters
686 Expression of Midgut ALP and ABCC Genes and Causes Resistance to *Bacillus thuringiensis*
687 Cry1Ac Toxin in Diamondback Moth. *PLoS Genet* 11(4): e1005124.
688 https://doi.org/10.1371/journal.pgen.1005124

689 Guzov VM, Unnithan GC, Chernogolov AA, Feyereisen R (1998) CYP12A1, a Mitochondrial
690 Cytochrome P450 from the House Fly. *Arc Biochem Biophys* 359(2):231-240.

691 Hartstack AW, Witz JA, Buck DR (1979) Moth traps for the tobacco budworm. *J Econ Entomol* 72:
692 519-522.

693 Hartstack AW, Lopez JD, Muller RA, Sterling WL, King EG (1982) Evidence of long range migration
694 of *Heliothis zea* (Boddie) into Texas and Arkansas. *Southwestern Entomol* 7:188–201.

695 Heckel DG (2020) How do toxins from Bacillus thuringiensis kill insects? An evolutionary
696 perspective. *Arc Insect Biochem Physiol* 104:e21673

697 Howe GA, Herde M (2015) Interaction of plant defense compounds with the insect gut: new insights
698 from genomic and molecular analyses. *Curr Opin Insect Sci* 9: 62-68.

699 Hudson RR (2002) Generating samples under a Wright–Fisher neutral model of genetic variation.
700 *Bioinformatics* 18(2): 337–338.

701 Hutchison, W. D., E. C. Burkness, P. D. Mitchell, R. D. Moon, T. W. Leslie, S. J. Fleischer, M.
702 Abrahamson, K. L. Hamilton, K. L. Steffey, M. E. Gray, et al. (2010) Areawide suppression of
703 European corn borer with Bt maize reaps savings to non-Bt maize growers. *Science* 330: 222–225.

704 Jin L, Wang J, Guan F, Zhang J, Yu S, Liu S, Xue Y, Li L, Wu S, Wang X, Yang Y, Abdelgaffar H,
705 Jurat-Fuentes JL, Tabashnik B, Wu Y (2018) Dominant point mutation in a tetraspanin gene
706 associated with field-evolved resistance of cotton bollworm to transgenic Bt cotton. *Proc Natl*
707 *Acad Sci USA* 115(46): 11760-11765.

708 Jurat-Fuentes JL, Adang MJ (2004) Characterization of a Cry1Ac-receptor alkaline phosphatase in
709 susceptible and resistant *Heliothis virescens* larvae. *Eur J Biochem* 271(15): 3127-3135.

710 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new
711 method for the analysis of genetically structured populations. *BMC Genet* 11:94. doi:10.1186/1471-
712 2156-11-94.

713 Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data.
714 *Bioinformatics* 27(21): 3070-3071. https://doi:10.1093/bioinformatics/btr521

715 Jones CM, Parry H, Tay WT, Reynolds DR, Chapman JW (2019) Movement ecology of pest
716 *Helicoverpa*: implications for ongoing spread. *Annu Rev Entomol* 64: 277–295.

717 Joußen N, Agnolet S, Lorenz S, et al. (2012) Resistance of Australian *Helicoverpa armigera* to
718 fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc Natl Acad Sci U S A*
719 109(38):15206-15211. doi:10.1073/pnas.1202047109

720 Kaur G, Guo J, Brown S, Head GP, Price PA, *et al.* (2019) Field-evolved resistance of *Helicoverpa zea*
721 (Boddie) to transgenic maize expressing pyramided Cry1A.105/Cry2Ab2 proteins in northeast
722 Louisiana. *J Inv Pathol* 163: 11-20.

723 Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357-359.

724 Laurentino TG, Moser D, Roesti M, Ammann M, Frey A, Ronco F, Kueng B, Berner D (2020)
725 Genomic release-recapture experiment in the wild reveals within-generation polygenic selection in
726 stickleback fish. *Nat Comm Biol* 11(1) DOI: 10.1038/s41467-020-15657-3

727 Leon RG, Dunnea JC, Gould F (2020) The role of population and quantitative genetics in modern
728 sequencing technologies to understand evolved herbicide resistance and weed fitness. *Pest Manag*
729 *Sci* DOI 10.1002/ps.5988

730 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000
731 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and
732 SAMtools. *Bioinformatics* 25(16):2078-9. doi: 10.1093/bioinformatics/btp352. Epub 2009 Jun 8.

733 Liu C, Xiao Y, Li X, Oppert B, Tabashnik BE, Wu K (2015) *Cis*-mediated down-regulation of a
734 trypsin gene associated with Bt resistance in cotton bollworm. *Sci Rep* 4: 7219.

735 Luttrell RG, Jackson RE (2012) *Helicoverpa zea* and Bt cotton in the United States. *GM Crops Food* 3:
736 213.

737 Magoc T, Salzberg S (2011). FLASH: Fast length adjustment of short reads to improve genome
738 assemblies. *Bioinformatics* 27(21): 2957-63.

739 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.
740 *EMBnet. journal*, *17*(1), 10-12.

741 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, *et al.* (2010). The Genome
742 Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.
743 *Genome Res* 20(9): 1297–1303. doi:10.1101/gr.107524.110

744 Melander AL (1914). Can insects become resistant to sprays? *J Econ Entomol* 7:167-173.

745 Micinski S, Blouin DC, Waltman WF, Cookson C (2008) Abundance of *Helicoverpa zea* and *Heliothis*
746 *virescens* in pheromone traps during the past twenty years in Northwestern Louisiana. *Southwest*
747 *Entomol* 33:139 - 49. http://dx.doi.org/10.3958/0147-1724-33.2.139

748 Morin S, Henderson S, Fabrick JA, *et al.* (2004) DNA-based detection of Bt resistance alleles in pink
749 bollworm. *Insect Biochem Mol Biol* 34(11):1225-1233. doi:10.1016/j.ibmb.2004.08.003

750 Morin S, Biggs RW, Sisterson MS, Shriver L, Ellers-Kirk C, Higginson D, Holley D, Gahan LJ,
751 Heckel DG, Carrière Y, Dennehy TJ (2003) Three cadherin alleles associated with resistance to
752 *Bacillus thuringiensis* in pink bollworm. *Proc Nat Acad Sci U S A* 100(9): 5004-9.

753 National Academies of Sciences, Engineering, and Medicine [NASEM] (2016) Genetically Engineered
754 Crops: Experiences and Prospects. Washington, DC: The National Academies Press. doi:
755 10.17226/23395.

756 Newcomb RD, Campbell PM, Ollis DL, Heah EC, Russell RJ, and Oakeshott JG (1997) A single
757 amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers
758 insecticide resistance on a blowfly. *Proc Natl Acad Sci USA* 94:7464 –7468.

759 Nosil P, Villoutreix R, de Carvalho CF, Farkas TE, Soria-Carrasco V, Feder JL, Crespi BJ, Gompert Z
760 (2018) Natural selection and the predictability of evolution in Timema stick insects. *Sci* 359
761 (6377):765-770.

762 Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Sci* 293(5536):1786-1790.

763 Pearce SL, Clarke DF, East PD *et al.* (2017) Genomic innovations, transcriptional plasticity and gene
764 loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa*
765 pest species. *BMC Biol* 15: 63. doi:10.1186/s12915-017-0402-6

766 Pélissié B, Crossley MS, Cohen ZP, and Schoville SD (2018) Rapid evolution in insect pests: the
767 importance of space and time in population genomics studies. *Curr Opin Insect Sci* 26: 8-16.

768 Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of genetic
769 differentiation and structure of mixed-ploidy level populations. *Mol Ecol Res* 13(5): 946-952.

770 Perera OP and Blanco CA (2011). Microsatellite variation in *Helicoverpa zea* (Boddie) populations in
771 the Southern United States. *Southwestern Entomol* 36(3):271-286.

772 R Development Core Team (2008). R: A language and environment for statistical computing. R
773 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-
774 project.org.

775 Reay-Jones FPF (2019) Pest Status and Management of Corn Earworm (Lepidoptera: Noctuidae) in
776 Field Corn in the United States. *J Int Pest Manag* 10(1):19 https://doi.org/10.1093/jipm/pmz017

Reisig DD, Huseth AS, Bacheler JS, Aghaee MA, Braswell L, *et al*. (2018) Long-term empirical and observational evidence of practical *Helicoverpa zea* resistance to cotton with pyramided Bt toxins. *J Econ Entomol* 111(4): 1824-33.

Reznick DN, Losos JB, Travis J (2019) From low to high gear: There has been a paradigm shift in our understanding of evolution. *Ecol Lett* 22:233–44.

Roush RT, McKenzie JA (1987) Ecological genetics of insecticide and acaricide resistance. *Ann Rev Entomol* 32:361-80.

Rubin CJ, Zody MC, Eriksson J, Meadows JR, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464(7288): 587.

Russell RJ, Scott C, Jackson CJ, et al. (2011) The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evol App* 4(2):225-248. DOI: 10.1111/j.1752-4571.2010.00175.x.

Seymour M, Perera OP, Fescemyer HW, Jackson RE, Fleischer SJ, Abel CA (2016) Peripheral genetic structure of *Helicoverpa zea* indicates asymmetrical panmixia. *Ecol Evol* 6(10):3198-207. https://doi:10.1002/ece3.2106

Tabashnik BE, Gould F, Carrière Y (2004) Delaying evolution of insect resistance to transgenic crops by decreasing dominance and heritability. *J Evol Biol* 17:904-912.

Tabashnik BE, Carrière Y (2017) Surge in insect resistance to transgenic crops and prospects for sustainability. *Nature Biotechnol* 35(10):926.

Tay WT, Mahon RJ, Heckel DG, Walsh TK, Downes S, James WJ, Lee SF, Reineke A, Williams AK, Gordon KHJ (2015) Insect resistance to *Bacillus thuringiensis* toxin Cry2Ab is conferred by mutations in an ABC transporter subfamily A protein. *PLoS Genet* 11: e1005534.

Taylor MF, Heckel DG, Brown TM, Kreitman ME, Black B (1993) Linkage of pyrethroid insecticide resistance to a sodium channel locus in the tobacco budworm. *Insect Biochem Mol Biol* 23(7): 763-775.

Tibshirani RJ, Efron B (1993). An introduction to the bootstrap. *Mono Statist Applied Prob* 57: 1-436.

U.S. Environmental Protection Agency [US-EPA] (1998) The environmental protection agency's white paper on Bt plant-pesticide resistance management. Washington: http://www.epa.gov/scipoly/sap/meetings/1998/february/finalfeb.pdf

U.S. Environmental Protection Agency [US-EPA]. (2018) Resistance of Lepidopteran Pests to Bacillus thuringiensis (Bt) Plant Incorporated Protectants (PIPs) in The United States. SAP Minutes No. 2018-06. FIFRA Scientific Advisory Panel Meeting, 17-19 July 2018 Arlington, VA.

Van Leeuwen T, Dermauw W, Mavridis K, Vontas J (2020) Significance and interpretation of molecular diagnostics for insecticide resistance management of agricultural pests. *Curr Opin Insect Sci* 39:69-76.

Ward CW (1976) Properties of the major carboxypeptidase in the larvae of the webbing clothes moth *Tineolla bisselliella*. *Biochim Biophys Acta* 429: 564–572.

815 Weetman D, Wilding CS, Neaffsey DE, Muller P, Ochomo E, Isaacs AT, *et al.* (2018) Candidate-gene
816 based GWAS identifies reproducible DNA markers for metabolic pyrethroid resistance from
817 standing genetic variation in East African *Anopheles gambiae*. *Sci Rep* 8:2920.

818 Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population-Structure.
819 *Evolution* 38: 1358-1370.

820 Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation:
821 Inference of a neutral model through trimming the distribution of FST. *Am Nat* 186(S1): S24-S36.

822 WuY (2014) Detection and mechanisms of resistance evolved in insects to Cry toxins from *Bacillus*
823 *thuringiensis*. *Adv Insect Physiol* 47, 297–342.

824 Yang F, González JC, Williams J, Cook DC, Gilreath RT (2019) Occurrence and ear damage of
825 *Helicoverpa zea* on transgenic *Bacillus thuringiensis* maize in the field in Texas, US and its
826 susceptibility to Vip3A protein. *Toxins* 11(2): 102.

827 Yang F, Head GP, Price PA, Gonzáleza JCS, Kerns DL (2020) Inheritance of *Bacillus thuringiensis*
828 Cry2Ab2 protein resistance in *Helicoverpa zea* (Lepidoptera: Noctuidae). *Pest Manag Sci* DOI
829 10.1002/ps.5916

830 Zhang H, Wu S, Yang Y, Tabashnik BE, Wu Y (2012) Non-Recessive Bt Toxin Resistance Conferred
831 by an Intracellular Cadherin Mutation in Field-Selected Populations of Cotton Bollworm. *PLoS*
832 *ONE* 7(12): e53418. https://doi.org/10.1371/journal.pone.0053418

833 Zhang H, Tang M, Yang F, Yang Y, Wu Y (2013) DNA-based screening for an intracellular cadherin
834 mutation conferring non-recessive Cry1Ac resistance in field populations of *Helicoverpa armigera.*
835 *Pesticide Biochem Physiol* 107(1):148-152.

836 Zhang Z, Teng X, Li F (2017) Knockdown of two Cadherin genes confers resistance to Cry2A and
837 Cry1C in *Chilo suppressalis.* *Sci Rep* 7: 5992.

838 Zhang M, Wei J, Ni X, Zhang J, Jurat-Fuentes JL *et al*. (2019) Decreased Cry1Ac activation by midgut
839 proteases associated with Cry1Ac resistance in *Helicoverpa zea*. *Pest Manag Sci* 75: 1099–1106.
840 https://doi.org/10.1002/ps.5224

841 **Tables**

842

843 **Table 1** - Candidate genes previously documented as being involved in Bt resistance in lab or field-
844 collected Lepidopteran species, their best alignments to the *H. zea* genome. We calculated the average
845 Weir and Cockerham's F_{ST} value for SNPs within the gene-containing region of each scaffold. The
846 number of SNPs used to calculate the average F_{ST} value can be found in parentheses. Whole genome
847 sequencing samples from the years 2002 and 2017 were used for the comparison. We also report
848 where the nearest SNP marker produced by ddRAD sequencing was found relative to each candidate
849 gene and its F_{ST} value. When ddRAD SNPs were found within the gene, the average F_{ST} value across
850 SNPs within the gene are reported. When the ddRAD SNP was found outside of the gene, the F_{ST}
851 value for the nearest SNP is reported.

852

Candidate Gene	**Uniprot/ Genbank ID**	**Citation**	***H. zea* scaffold**	**Gene Position (bp)**	**Avg F_{ST} (Num SNPs)**	**Nearest ddRAD SNP (bp)**	**ddRAD Avg F_{ST} (Num SNPs)**
alp	ACP39715.1	Jurat-Fuentes and Adang 2004	KZ117832.1	1,946 - 16,347	-0.011 (171)	NA	NA
apn1	Q11000.1	Gill *et al.* 1995	KZ118301.1	140,429 - 161,526	-0.005 (353)	in gene	-0.003 (3)
apn4	AF378666.1	Banks *et al.* 2001	KZ118301.1	168,017 - 174,495	0.056 (170)	22,238	-0.005 (1)
abcA2	KP259911.1	Tay *et al.* 2015	KZ118207.1	171,103 - 192,554	0.002 (217)	in gene	0.006 (2)

abcC2	GQ332571.1	Gahan *et al.* 2010	KZ118297.1	98,882 - 114,139	-0.004 (100)	3,178	-0.004 (1)
abcG1 (white)	KM667971.1	Guo *et al.* 2015a	KZ118424.1	2,680 - 24,508	0.015 (58)	12,964	0.001 (1)
calp	AAK85198.1	Gahan *et al.* 2001	KZ117563.1	271,234 - 288,320	0.015 (201)	56,926	-0.007 (1)
cad2	AGG36450.2	Zhang *et al.* 2017	KZ118195.1	87,164 - 95,177	-0.003 (44)	20,535	-0.003 (1)
cad-86C	*H. zea* assembly v. 1.0 ID 537580	Fritz *et al*. 2020	KZ117463.1	550,935-603,715	-0.008 (1079)	in gene	-0.004 (2)
map4K4	AKC96184.1	Guo *et al.* 2015b	KZ118817.1	92,551 - 104,845	-0.018 (65)	35,125	0.005 (1)
tspan1	AYM54687.1	Jin *et al*. 2018	KZ118424.1	106,132 - 117,025	-0.008 (164)	27,211	0.003 (1)

853

854 **Table 2** - Putative amino acid substitutions (AA Sub) for two *carboxypeptidase Q* genes (*cpq1* and
855 *cpq2*) and *cyp333b3*. We quantified frequency changes in the codon present in the reference genome
856 (Ref Codon) for *H. zea* populations collected in 1998 (n chr = 40-52), 2002 (n chr = 26), 2012 (n chr =
857 22), and 2017 (n chr = 22). A Freeman-Halton extension of a Fisher's exact test with a bonferroni
858 correction (α = 0.006) determined statistical significance (*). *Cpq1* extended from 8583 to 10,070 bp
859 on scaffold 569, and *cpq2* extended from 151,736 to 160,706 bp, while cyp333b3 extended from
860 32,819 to 39,355 bp on scaffold 1612. All three genes are transcribed from the reverse strand.
861

Scaffold	**Gene**	**SNP Pos (bp)**	**AA Sub**	**Ref Codon**	**2002 freq**	**2012 freq**	**2017 freq**	**p-val**
KZ118395.1	*cpq1*	9,409	A → V	GCT	0.73	0.32	0.14	0.0001*
KZ118395.1	*cpq2*	152,183	T → I	ATA	0.83	0.73	1	0.022
KZ118395.1	*cpq2*	152,595	A → V	GTA	0.84	0.77	1	0.049
KZ117131.1	*cyp333b3*	35,685	S → T	AGC	0.65	0.32	0.18	0.0001*
KZ117131.1	*cyp333b3*	36,339	K → N	TTT	0.69	0.27	0.09	0.0003*
KZ117131.1	*cyp333b3*	36,579	G → V	GGA	0.69	0.27	0.09	< 0.0001*
KZ117131.1	*cyp333b3*	36,959	T → A	GCC	0.62	0.95	1	< 0.0001*
KZ117131.1	*cyp333b3*	37,490	S → N	AGC	0.73	0.32	0.09	0.003*

862

863
864 **Figures**
865

866
867 **Figure 1 -** Posterior membership probabilities for *H. zea* individuals collected in Bossier Parish, LA, in 2002, 2007, 2012, and 2016.
868 Genotypic clustering used ddRAD-seq generated SNP data and assumed a prior number of clusters (k) equal to two.
869

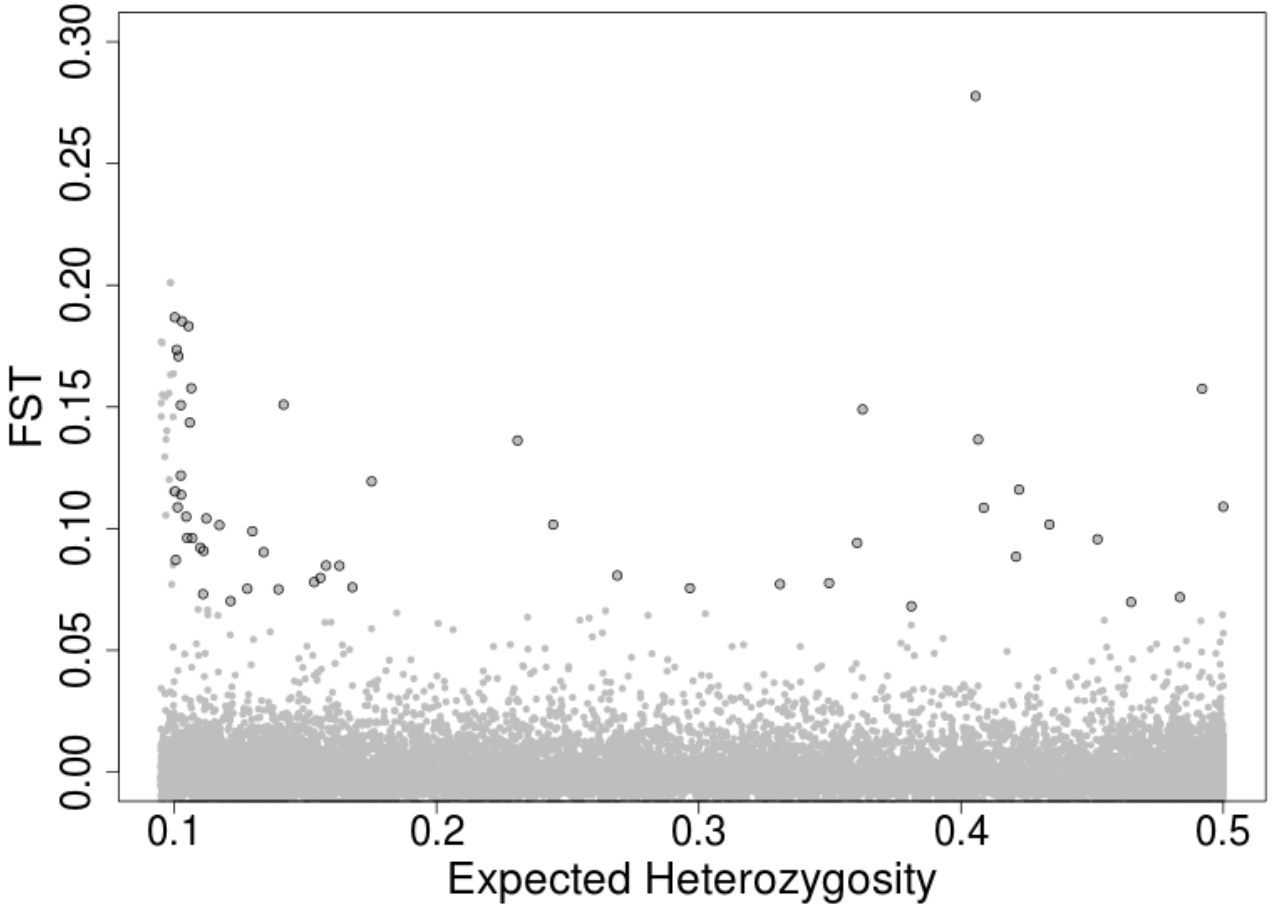

Figure 2 - F_{ST} outlier analysis of *H. zea* collected from Bossier Parish, LA in the years 2002, 2007, 2012, and 2016. Each grey point represents one of the 14,398 SNPs examined in our total dataset. Those circled in black represent SNPs that were identified as outliers and indicate they were under directional selection.

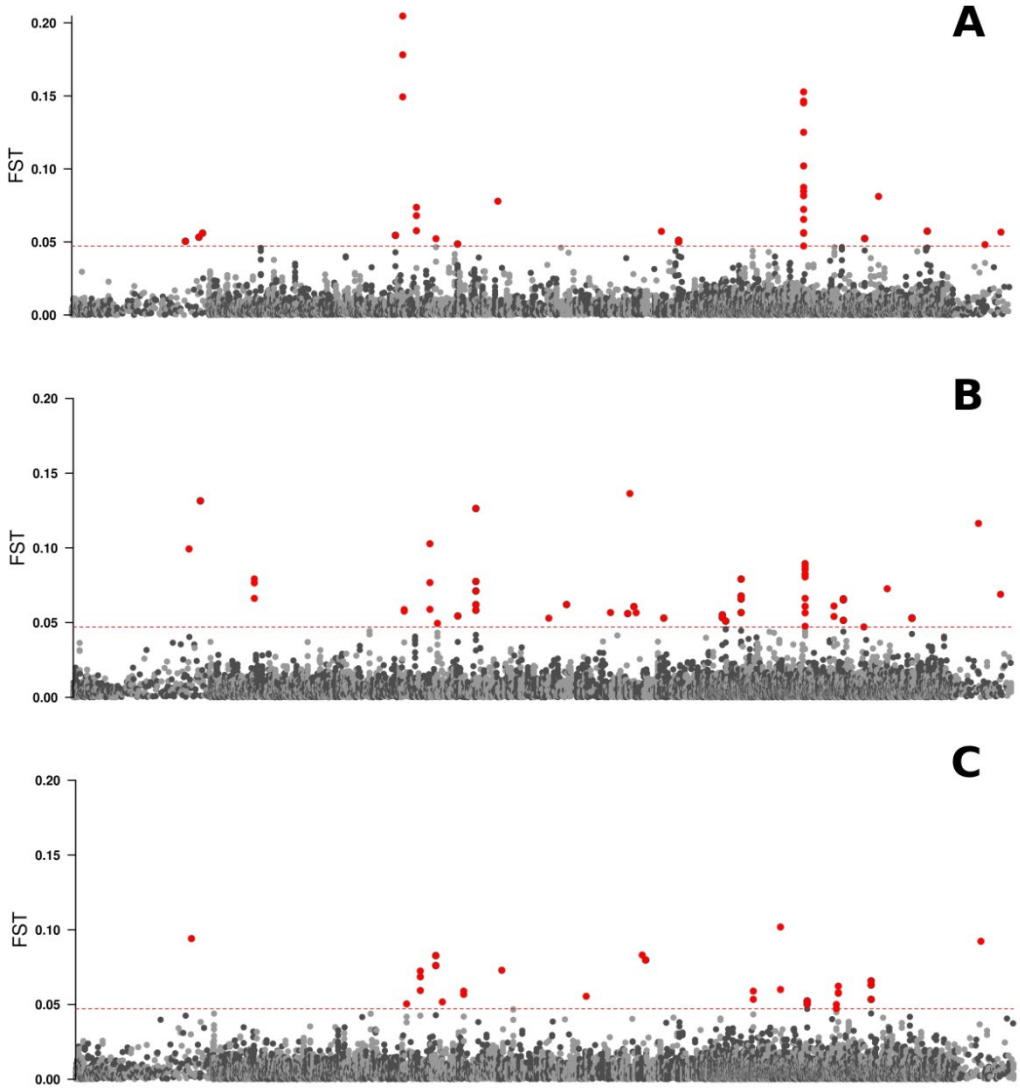

Figure 3 - Manhattan plots of genomic divergence for individuals collected from pheromone-baited traps in Bossier Parish, LA in A) 2002 and 2017, B) 2002 and 2012, and C) 2012 and 2017. Alternating light and dark grey points demarcate averaged F_{ST} values for 40kb windows with a 10kb step for each of the 2,975 genomic scaffolds. Red points represent statistically significant genetic divergence.

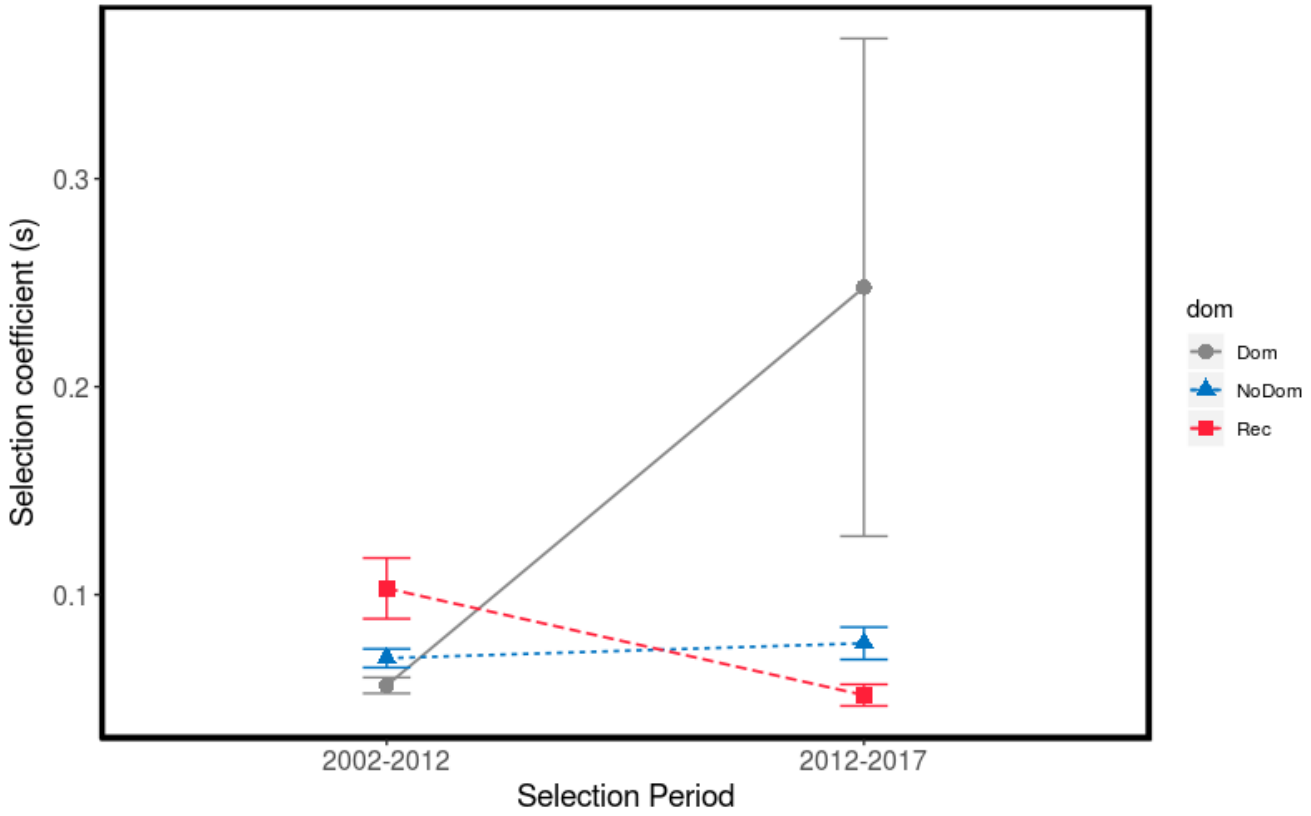

Figure 4 - Values of *s*, the selection coefficient, required to produce the observed allele frequency changes at SNPs within protein-coding regions in field-collected *H. zea* during two selection periods: 2002-2012 and 2012-2017. Points represent the mean values (n = 6 SNPs), and error bars represent standard error of the mean.

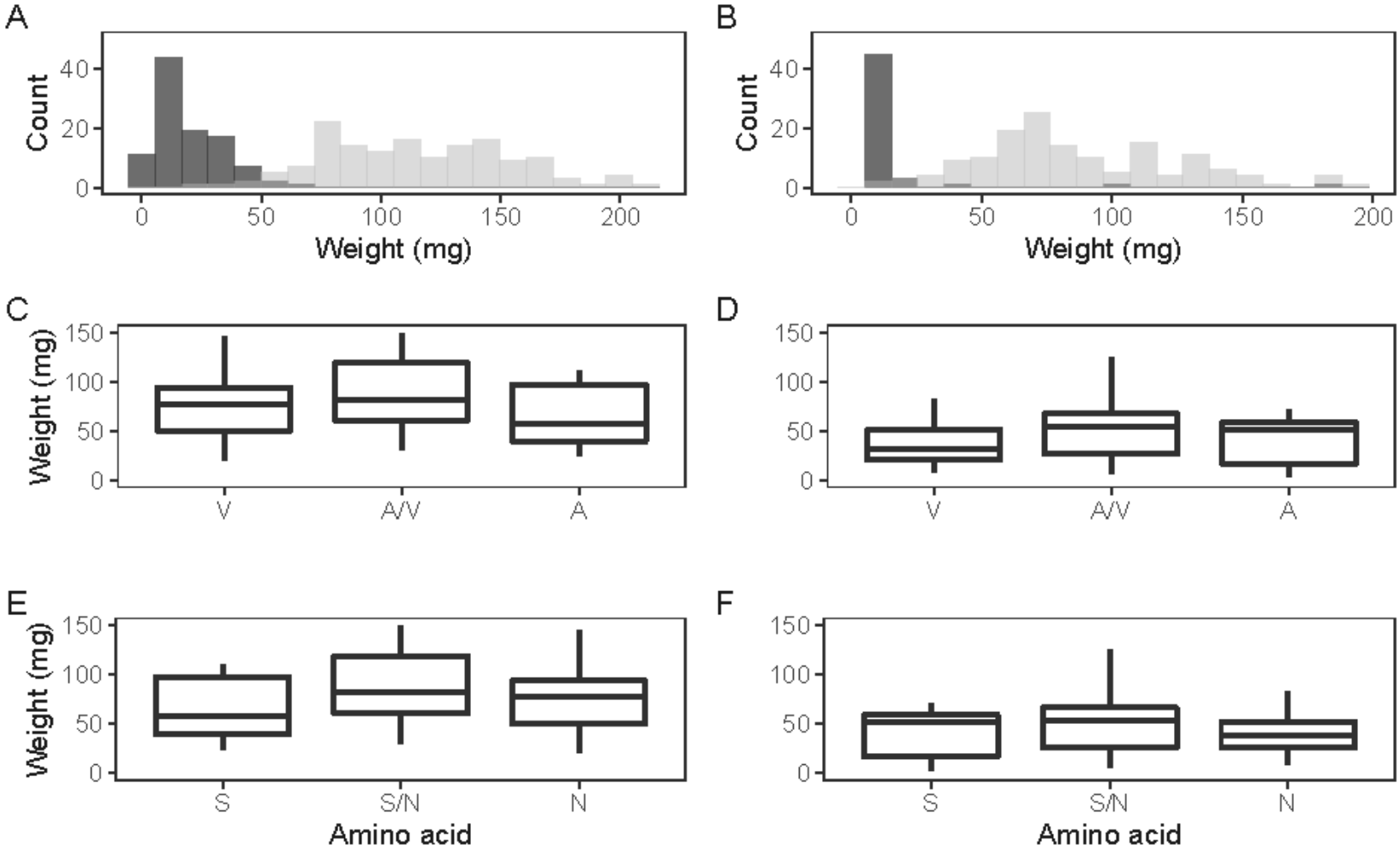

Figure 5 – Association of F_2 larval growth phenotypes on Cry-treated diet with nonsynonymous substitutions that cause protein coding changes in *cpq1* and *cyp333b3*. After one week of growth with **A.** Cry1 or **B.** Cry1 + Cry2 containing diet the susceptible control strain (light gray) grew 7 and 10 times less on the two diets respectively than the resistant Maryland population (dark gray). No significant association was detected between Cry toxin resistance phenotype and non-synonymous substitutions in the sweep region for F_2 progeny of these lines according to a Kruskall Wallis test. Box plots compare weight after one week of growth on Cry1 (**C, E)** or Cry1 + Cry2 containing diet (**D, F)** for F_2 progeny across genotypes in two genes, *cpq1* (**C, D**), *cyp333b3* (**E, F**).