# Genome resequencing reveals rapid, repeated evolution in the Colorado potato beetle, *Leptinotarsa decemlineata*

Benjamin Pélissié[1,2], Yolanda H. Chen[3], Zachary P. Cohen[1], Michael S. Crossley[1,4], David J. Hawthorne[5], Victor Izzo[3], Sean D. Schoville[1*]

[1] Department of Entomology, University of Wisconsin-Madison, Madison, WI 53706, USA

[2] Current address: Department of Biology, University of Nebraska-Kearney, Kearney, NE 98849, USA

[3] Department of Plant and Soil Science, University of Vermont, Burlington, VT 05405, USA

[4] Current address: Department of Entomology, University of Georgia, Athens, GA 30602, USA

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

* Corresponding author: Sean Schoville, Dept. Entomology, UW-Madison, 1630 Linden Drive, 637 Russell Labs, Madison WI 53706 Email: sean.schoville@wisc.edu

## Abstract

**Background:** Insecticide resistance and rapid pest evolution threatens food security and the development of sustainable agricultural practices. An improved understanding of the evolutionary mechanisms that allow pests to rapidly adapt to novel control tactics will help prevent economically damaging outbreaks. The Colorado potato beetle (CPB), *Leptinotarsa decemlineata*, is a global super-pest that rapidly evolves resistance to insecticides. Using whole genome resequencing and transcriptomic data focused on its ancestral and pest range in North America, we assess evidence for three, non-mutually exclusive models of rapid evolution: pervasive selection on novel mutations, rapid regulatory evolution, and repeated selection on standing genetic variation.

**Results:** Population genomic analysis demonstrates that CPB is geographically structured, even among recently established pest populations. Pest populations exhibit only modest reductions in nucleotide diversity, relative to non-pest ancestral populations, and show evidence of recent demographic expansion. Genome scans of selection provide clear signatures of repeated adaptation across different CPB populations, with especially strong evidence that insecticide resistance involves selection of different genes in different populations. Similarly, analyses of gene expression show that constitutive upregulation of candidate insecticide resistance genes drives distinctive population patterns.

**Conclusion:** CPB evolves insecticide resistance repeatedly across agricultural regions, and oftentimes at the same loci, supporting a prominent role of polygenic evolution from standing

genetic variation. Despite expectations, we do not find support for strong selection on novel mutations, or rapid evolution from selection on regulatory genes. An important future goal will be to understand how polygenic resistance phenotypes spread among local pest populations, in order to refine integrated pest management practices to maintain the efficacy and sustainability of novel control techniques.

**Key Words:** population genomics; insecticide resistance; rapid evolution; genetic adaptation; regulatory evolution; insect pest; polygenic adaptation

## Background

Herbivorous pests cause an estimated 18-20% damage to crops and cost nearly $470 billion annually on a global scale [1]. The ability of insect pests to evolve resistance to insecticides threatens food security and the development of sustainable agricultural practices, especially when their rate of evolution outstrips the development of novel control strategies [2-4]. This is the case with insect 'super-pests,' which repeatedly evolve insecticide resistance even as they are faced with completely novel insecticides, thus perpetuating the arms race that defines the pesticide treadmill [5]. Curiously, particular super-pest species or even select populations are more likely to adapt to new compounds, suggesting that there is a genetic basis in the propensity to evolve resistance [6]. Yet, despite more than 60 years of research on the evolution of resistance [7], the relative importance of alternative mechanisms that underlie the evolutionary potential for pesticide resistance evolution are still unclear [8, 9]. Although a considerable effort has been placed on understanding the proximal molecular control of resistance [10], broader questions about the genetic complexity of resistance, mode of selection and geographical extent of adaptation have rarely been studied [11-13]. While population genetic models of resistance management have been highly effective in certain management scenarios [14], observed patterns of insecticide resistance evolution defy many of the assumptions of our evolutionary models [15]. Recent genomic resequencing datasets suggest that resistance evolution is sometimes geographically and genetically complex [16-19].

A key goal should be to understand the evolutionary processes that allow species to become pests, particularly the mechanisms underlying phenotypic shifts that result in economically damaging pest outbreaks [8, 20]. A prevalent view is that pests, including invasive species, often retain substantial genetic diversity that facilitates evolution in agroecosystems [21-

23]. However, there is increasing recognition that evolution can be rapid irrespective of levels of standing genetic diversity [24]. While not all pests exhibit rapid rates of adaptation to insecticides [6], insect super-pests often demonstrate repeated rapid evolution [25, 26]. Rapid evolution is defined as a shift in phenotype from underlying variation in exceptionally few generations [27], and can occur as a result of several mechanisms [28]. First, selection can act on novel mutations, which may arise frequently if pests have intrinsically large population sizes (much greater than $>>10^6$) and are not mutation-limited [29, 30]. This could lead to repeated evolution of resistance among different populations, most likely with independent mutations at different loci underlying resistance phenotypes. Second, as a special case of the first mechanism, key mutational changes could affect a master regulatory gene [31-33], where mutations drive expression of the same downstream molecular pathways in different populations. Rapid gene regulatory evolution has been raised as a possible mechanism underlying repeated evolution of pesticide resistance in the spider mite *Tetranychus urticae*, where it has been linked to a transcriptional cascade in xenobiotic detoxification [34]. Third, an alternative pathway of rapid evolution would draw on standing genetic variation [7]. Standing genetic variation is increasingly viewed as a common source of rapid adaptive variation [35], because the initial frequency of mutations in a population determine the rate at which populations respond to selection pressures [36, 37]. While population size must typically be large to retain large reservoirs of standing variation, admixture among divergent populations can increase standing variation [38, 39]. Furthermore, standing genetic variation can also be present in the form of redundancy in molecular pathways that are critical to pesticide resistance phenotypes [e.g. in generalist herbivorous insects that specialize on toxic plants: 40, 41, 42], rather than allelic diversity *per se*. It should be emphasized that these mechanisms of adaptation need not be

exclusive, yet it remains unclear how each contributes to the evolutionary success of the top arthropod super-pests. Emerging genomic datasets provide the opportunity to detect and quantify the importance of different mechanisms underlying rapid evolutionary change by screening for genomic signatures of selection [8].

The Colorado potato beetle (CPB), *Leptinotarsa decemlineata,* is a global super-pest and an especially tractable exemplar of rapid evolution to insecticides. CPB has evolved resistance to over 50 different insecticides in all the major classes, in some cases within the first year of use [43]. CPB has demonstrated an ability to rapidly evolve in response to a wide range of environmental pressures, including host-plant defenses and climatic variability [44, 45]. This super-pest originated in the Great Plains region of the U.S. [46], following a host shift to potato (an introduced crop) in the mid-19th century (around 1866) that allowed for rapid spatial expansion from Nebraska to the Eastern U.S. in a 20 year period and colonization of Eurasia by the early 1900s [47-49]. Despite rapid spatial expansion, populations are genetically differentiated [49, 50] and insecticide resistance is geographically heterogeneous [51], even over local landscape scales [52]. In particular, beetles from Long Island, New York are known to have the highest levels of baseline resistance and are typically the first populations to develop resistance to all compounds [53], while populations in the Pacific Northwest remain susceptible to insecticides despite an equivalent duration of usage and comparable treatment practices [54, 55]. Non-pest populations are found in the Great Plains and Mexico, where they use ancestral host plants (primarily *Solanum rostratum*) [56]. Closely related congeners in the genus *Leptinotarsa* are sympatric in the southern part of CPB's geographical range [57]. By integrating across this diversity, CPB can serve as a model for understanding evolutionary mechanisms that facilitate and constrain rapid evolution.

Here we leverage the recent publication of the CPB genome [58] to investigate whether repeatable patterns of evolution occur in highly resistant pest populations. We compare CPB genomic and transcriptomic variation across populations in the U.S., Mexico, and Europe, as well as closely related *Leptinotarsa* species, to assess three competing models of rapid evolution: pervasive selection on *de novo* mutation (independent hard selective sweeps in geographically separate populations), rapid regulatory evolution, or repeated selection on standing genetic variation (see **Table 1** for predictions). We also provide detailed description of genomic diversity patterns, evolutionary relationships, and the population history of CPB pest lineages, in order to understand how expansion history has influenced geographical variation in insecticide resistance. Over the long-term, by improving our understanding of the evolutionary processes and genomic mechanisms underlying the ability to repeatedly evolve insecticide resistance in super-pests, integrated pest management strategies can be developed to provide more sustainable agricultural practices [2, 22].

## Results

### *Extensive Genomic Diversity within CPB*

We examined short-read whole genome sequences for a geographically dispersed set of 85 samples, including six geographically proximate pairs of susceptible and resistant samples, as well as nine additional *Leptinotarsa* species (**Fig. 1** and **Additional File 1: Table S1** and **Fig. S2**). Employing best practices in genotype ascertainment, we sequenced each sample (most resulted in coverage >4x, **Fig. S3**), and found that CPB shows considerable genomic diversity, with 76,647,868 single nucleotide polymorphisms (SNPs) recovered from the nuclear genome (**Additional File 1: Table S2**; the CPB genome is estimated to be ~670 Mb in size). Following

variant recalibration, there were a total of 47,969,460 SNPs, of which 30,973,249 were in intergenic regions.

We estimated genome-wide nucleotide diversity ($\pi$) using a 10 Kb sliding window. Within-population genetic diversity of non-pest CPB samples from the inferred ancestral source population, the U.S. Plains region, was high with an average $\pi = 0.005$. Comparison of nucleotide diversity between the Plains and U.S. pest populations shows a clear reduction in nucleotide diversity (average $\pi = 0.0028$ and 0.003 for co-located pairs of pesticide resistant and susceptible populations, respectively; **Additional File 1: Fig. S4**). However, pooling co-located pairs of samples increased nucleotide diversity in a given agricultural region by 40% and eroded the difference from the Plains samples, showing that high levels of variation have been retained in agricultural regions. Individual heterozygosity appeared to be reduced relative to expectations under random mating, as measured by the inbreeding coefficient $F_{IS}$. Inbreeding was higher in pest populations (on average $F_{IS} = 0.603$ and 0.557 for susceptible and resistant populations, respectively; **Fig. S5**), relative to Plains individuals (on average $F_{IS} = 0.531$). The New Jersey lab population, which was maintained as a breeding colony for pesticide assays, showed the highest level of inbreeding ($F_{IS} = 0.723$). Susceptible and resistant pest samples had a comparable number of private alleles (93,407 vs. 89,599, respectively; **Additional File 1: Fig. S6** and **S7**), but Plains samples had three times as many private alleles (262,140 vs. 91,503, respectively). Combining observations of nucleotide diversity, inbreeding, and private alleles suggests that pest lineages lost genetic variation as they expanded into agricultural habitats, but do not appear to have suffered from a strong genetic bottleneck.

***Evolutionary Diversification and Demographic History***

*Leptinotarsa decemlineata* population structure is clearly driven by geographic isolation. Phylogenetic reconstruction (**Fig. 1A**) clearly separated the Mexican samples from all U.S. samples. All CPB populations from potato growing regions in the U.S. and Eurasia formed a well-supported monophyletic group, while separate clades contained samples in the Pacific Northwest, Southwest U.S. and the U.S. Plains region. The Arizona CPB sample was found outside the CPB clade, next to the *L. lineolata* sample, suggesting high divergence and possibly cryptic species status. Interestingly, Mexican CPB and other *Leptinotarsa* species were mixed together, also suggesting that Mexican CPB belong to an unidentified cryptic *Leptinotarsa* species. Due to their significant genetic divergence, Mexican CPB, the Arizona sample, and other *Leptinotarsa* species were removed from downstream analyses of population structure and demographic change, as their inclusion would violate statistical assumptions in those approaches. Examining U.S. CPB samples, genetic divergence among populations was modest (average pairwise $F_{ST}$ = 0.09; **Additional File 1: Fig. S8** and **Table S3**), but exceeded >0.1 when New Jersey or the Michigan resistant population were compared. Principal component and population ancestry analyses converged in showing clear geographical patterns of population genetic structure (**Fig. 1B**; **Additional File 1: Figs. S9-S12**). Admixture-based clustering supported six populations, which represent the New-Jersey population, a western population (Oregon plus Idaho), a Plains population, a Midwestern population, a distinctive Michigan resistant population, and an Eastern U.S. population (which includes the introduced European samples). Admixture tests using the *D* statistic were examined among CPB pest and non-pest populations, as well as with other *Leptinotarsa* species. These tests provided limited evidence of admixture contributing to genetic diversity of pest lineages (**Fig. 2**), ruling out the hypothesis of standing genetic variation increasing in the pest lineage as a result of hybridization. The highest $D_{min}$

values suggest historical admixture from Mexico to the Plains (0.036) and Mexico to the Western populations (0.044), and limited ongoing gene flow between New York and the Michigan susceptible population ($D$ = 0.015). Similarly, an assessment of admixture with other *Leptinotarsa* species (**Additional File 1: Fig. S13**) does not suggest recent gene flow into pest populations.

Demographic reconstruction of CPB populations using SMC++ and Stairway plot analysis (**Fig. 1C** and **Additional File 1: Fig. S14-S15**) showed consistent population size fluctuations through time, with similar trajectories for all pest populations. Nearly all agricultural populations exhibited recent population size increases in the SMC++ analysis, most notably in the Eastern U.S. The split time analysis suggested an early split of the western populations from the Plains region, and subsequent near-simultaneous split times of Midwest and Eastern populations (**Additional File 1: Fig. S16)**. Applying a mutation rate from other insect taxa ($2.1 \times 10^{-9}$ substitutions per site per generation) suggests that populations contracted between 300k and 100k years ago, expanded between 200k and 70k years, and declined again until 10k to 5k years ago. The splits of most pest populations from the sampled populations in the Plains occurred between 21k and 11k years, during the transition from the late Pleistocene to early Holocene.

***Genome-wide Patterns of Natural Selection***

Genomic diversity across CPB's geographical range was scanned for evidence of natural selection by identifying outlier SNPs in comparisons of population differentiation and SNPs that were highly correlated to environmental predictor variables. Population differentiation tests identified 0.37% of all SNPs as outliers (i.e. 65,815 out of 17,599,906, with a false discovery rate, or FDR, of 0.01%). A total of ~32% of the outlier SNPs could be assigned to 8,760 known

genes (**Additional File 1**: **Table S4** and **Table S5**; gene list provided in **Additional File 2**). Of these genes, 336 were linked to candidate insecticide resistance genes, including 205 genes involved in detoxification pathways, 91 target-sites, and 40 genes involved in cuticular development. The well-known voltage-sensitive sodium channel gene (LDEC011942) that provides knockdown resistance to pyrethroids was included among the target-site genes. Based on a gene set enrichment analysis, over-represented gene ontology terms were linked to insecticide resistance and/or stress (**Additional File 1**: **Fig. S17**). For biological processes, GO terms included oxidation-reduction process and response to oxidative stress (and multiple nested terms), among others. For cellular components, terms linked to insecticide resistance included voltage-gated sodium channel and acetylcholine-gated channel complexes, presynaptic active zone and synapse, and integral component of the membrane. Among the molecular functions, terms such as heme and zinc ion binding (including iron ion binding), extracellular ligand-gated ion channel activity (including voltage-gated sodium channel and acetylcholine receptor activity), glutathione transferase activity, ABC transporter activity via the term ATPase activity coupled to transmembrane movement, and peroxidase activity (including CYP monooxygenase activity) were associated with insecticide resistance.

To test for patterns of natural selection driven by environmental factors, we also employed gene environment association analysis to examine five different, ecologically-relevant predictors of natural selection on the genome: latitude, elevation, precipitation, minimum temperature in the coldest month and potato land cover. Only 0.02% of the analyzed SNPs (4,098 out of 17,599,906 SNPs, with an FDR of 0.01%) were significantly associated with at least one environmental variable (**Additional File 1**: **Table S6** and **Fig. S18**; gene list provided in **Additional File 2**). A total of 67.6% of the SNPs were associated with precipitation, 15.5%

with latitude, 10.7% with potato land cover, 2.8% with elevation, and 3.7% with temperature. Of all significant SNPs, 29% were found in 816 known genes, including 42 resistance-related genes (28 involved in metabolic detoxification mechanisms, 3 in cuticle development, and 11 target-site genes; **Additional File 1**: **Table S7**). Based on a gene set enrichment analysis, over-represented gene ontology terms were associated with insecticide resistance and/or stress (**Additional File 1**: **Fig. S19**). Among the biological processes, terms included chemical synaptic transmission, oxidation-reduction process, proteolysis, defense response, and DNA repair. Among the cellular components, terms included synapse and presynaptic active zone, as well as integral component of the membrane. Among the molecular functions, terms included heme and iron ion binding, carboxylic ester hydrolase activity, extracellular ligand-gated ion channel activity, oxidoreductase activity (including monooxygenase activity), and ubiquitin binding.

The outlier-based and environmental-association genome scans leverage different models to detect selection, but a comparison of the results shows that a total of 557 genes (see **Additional File 2**) were shared in both tests, and 29 of these are candidate insecticide resistance genes (**Table 2**). The largest group represents xenobiotic detoxification genes, with nine ABC transporters, seven CYP genes, two esterase genes, one MFS gene and one GST gene represented. Target-site genes included four voltage-dependent channel genes and two nicotinic acetylcholine receptors, and three cuticle genes overlap in both tests. Gene set enrichment analysis of significant genes identified as overlapping in the two genome scan tests (**Additional File 1**: **Fig. S20**) showed enrichment of gene ontology terms associated with insecticide resistance and stress. Among biological processes, terms included chemical synaptic transmission, oxidation-reduction process, proteolysis, DNA repair, and chloride transport, transmembrane transport, and ion transport. Terms associated with cellular components included

integral component of membrane, synapse, and presynaptic active zone. Terms associated with molecular functions included pathways such as heme and iron ion binding, ATPase activity coupled to transmembrane movement, GABA and G-protein coupled receptor activity, and monooxygenase and oxidoreductase activity.

***Local Adaptation to Insecticides***

To examine the geographical occurrence of selection events, we employed a haplotype-based approach that examines shifts in haplotype frequency along branches of a population tree. Due to the fragmented nature of the reference genome, we examined haplotype frequencies on the longest 95 genomic scaffolds (encompassing ~21% of the genome, all >1 Mb). Our analyses show that 1.1% (72,386 SNPs), 0.01% (7,826 SNPs) and $0.16e^{-3}$% (1,106 SNPs) of the markers were significant at $\alpha = 0.01$, 0.001 and 0.0001, respectively (see list in **Additional File 2**). SNPs were grouped into regions, where each region was separated by at least 1 Kb up- and downstream. This resulted in 1,169 selection regions at $\alpha = 0.01$, 140 regions at $\alpha = 0.001$, and 24 regions at $\alpha = 0.0001$ (**Additional File 1**: **Table S8**). Excluding one extremely long region (35 Kb in length), the average length of the most significant haplotypes ($\alpha = 0.0001$) was 3.1 Kb; **Additional File 1: Fig. S21**). On average (across $\alpha$ levels and branch association thresholds), these regions in the genome were repeatedly selected in multiple populations (on average, in 6.06 branches of the population tree, and only 4.87% of these selection events were singularly associated with one population; **Additional File 1: Table S9**). These singular regions tended to be short (1.1 Kb on average). Out of all selected regions ($\alpha = 0.01$), 319 were found in 224 genes and only 3.8% were singular (**Additional File 1: Fig. S22**). Recalling that this only represents ~21% of the genome, 24 regions comprising 16 genes were candidate insecticide resistance-

277 associated genes, including six ABC transporters, two esterases, one olfactory receptor, one
278 nicotinic acetylcholine receptor, two genes associated with glutamate pathways and four growth
279 factors (**Additional File 1: Table S10**).

280 Most haplotype-based selection events in the candidate insecticide resistance genes (19
281 out of 24 regions) were less than 1 Kb long (**Additional File 1: Fig. S23**) and present (22 out of
282 24 regions) on nine to eleven branches of the population tree (**Fig. 3**; **Additional File 1: Fig.**
283 **S24**). Selection at these candidate regions (19 out of 24) were shared between Western and
284 Eastern lineages (**Additional File 1: Fig. S25**), which were the most genetically distinct and
285 geographically isolated populations. Furthermore, several of these candidate genes
286 (LDEC004355, LDEC005089, and LDEC002775) appeared to have multiple regions under
287 selection, with population-specific patterns. The observed patterns at insecticide resistance
288 candidates suggests that repeated selection on the same set of protein-coding genes is prevalent
289 among populations. Over 150 genes (68.5%) identified in the haplotype-based test were also
290 identified in the outlier test (**Additional File 1: Fig. S26**), including 11 of the 16 candidate
291 insecticide resistance genes. Seven of these were ABC transporters, and notably three were
292 significant in all selection tests (the ABC subfamily C/multidrug associated gene LDEC002518,
293 and two ABC subfamily G genes, LDEC002775 and LDEC005530). The other ABC genes
294 included: the subfamily F gene LDEC004565, and three additional multidrug associated genes
295 LDEC005089, LDEC003183, and LDEC002116. The remaining genes included one target site
296 gene, the acetylcholine receptor β subunit (LDEC002850), one cuticle protein (LDEC003397),
297 one transient receptor potential (TRP) gene (LDEC003216), and one odorant binding gene
298 (LDEC003898).

We examined gene expression profiles to test whether patterns of gene regulatory evolution were shared among geographical regions for a subset of previously published CPB population samples (**Additional File 1: Table S11**). For quality control, we first assessed whether differences in experimental design influenced the expression of candidate insecticide resistance genes (see detailed results in **Additional File 1)**. Based on these comparisons, we determined that regional population differences could be compared for adults from field populations irrespective of generation sampled, but lab reared larvae needed to be compared separately. For the larval comparison, we removed samples representing an insecticide induction treatment, focusing our analysis on constitutive differences in gene expression. These comparisons showed strong geographical differences in overall gene expression profiles (**Additional File 1: Fig. S27** and **Fig. S28**). Focusing on significantly differentially expressed candidate insecticide resistance genes, local populations showed divergent patterns of constitutive upregulation among populations (**Fig. 4** and **Additional File 1: Fig. S29**; see gene list in **Additional File 2**). Seven differentially expressed candidate insecticide resistance genes were found among the adults and 84 among larvae (**Additional File 1: Fig. S29** and **Table S12**), with only one esterase (LDEC019310) and one ABC transporter (LDEC004154) common to both datasets. A gene set enrichment analysis of the differentially expressed gene list in among-population comparisons of adults (**Additional File 1: Fig. S30**) showed enrichment of terms associated with insecticide detoxification and/or stress, including heme and ion iron binding, oxidoreductase activity and dioxygenase activity, proteolysis, transport, defense response, and integral component of the membrane. Among larvae (**Additional File 1: Fig. S31**), there was enrichment of terms associated with regulatory changes to gene networks underlying insecticide detoxification or stress, such as glutathione transferase activity, hexachlorocyclohexane

metabolism, oxidoreductase and monooxygenase activity, gap junction channel activity, proteolysis, substrate-specific transmembrane transporter activity, heme and iron ion binding, innate immune response, and integral component of the membrane.

Although 26 transcription factors were significantly differentially expressed among the four larval populations (**Additional File 1: Fig. S32** and **Table S13**), they are not known to be associated with detoxification pathways. These results do not support rapid evolution of regulatory genes, as different genes, and in some cases different molecular pathways, are favored in each regional population. Gene set enrichment analysis of the overlapping set of significant genes in larvae and adult differential expression analysis gene sets (**Additional File 1: Fig. S33**) showed shared regulatory changes in gene networks linked to insecticide detoxification and stress, such as oxidoreductase and dioxygenase activity, heme and iron ion binding, proteolysis and integral component of membrane. The shared enrichment of lipid metabolism might also be related to insecticide detoxification (through an interaction with oxidation-reduction or membrane-transport processes), rather than metabolism per se.

Finally, we examined a global intersection of gene ontology terms that were enriched in both the genome scans and differential expression datasets (**Additional File 1: Fig. S34**). Six gene ontology terms were identified, each of which has been related to insecticide resistance in prior studies: biological process terms include oxidation-reduction process and proteolysis, while cellular components include integral component of the membrane, and molecular functions include oxidoreductase activity acting on paired donors with incorporation or reduction of molecular oxygen, heme binding, and iron ion binding.

## Discussion

345 Population genomics is increasingly providing insight into the evolutionary mechanisms
346 that give rise to super-pests and holds promise for improving pest management practices [8]. Our
347 study provides the first comprehensive genomic and geographical assessment of genome-wide
348 patterns of genetic variation for a super-pest in the center of its origin. Whole genome variation
349 demonstrates that the Colorado potato beetle (CPB) is geographically structured, including
350 among pest populations, and corroborates evidence from microsatellite markers that CPB pest
351 populations are most closely related to populations in the Great Plains instead of Mexico [46, 59,
352 60]. In evaluating the competing (but not mutually-exclusive) mechanisms of rapid evolution,
353 our data suggest that different loci are selected across growing regions, and among different
354 populations within regions, with statistical tests providing a consistent pattern of repeated
355 evolution at many candidate insecticide resistance genes. Below we discuss in turn the evidence
356 that supports polygenic evolution from standing genetic variation, selection on *de novo* mutation,
357 and rapid regulatory evolution. We close by discussing the pest management consequences of
358 these modes of rapid evolution.
359
360 *Evidence for repeated selection on standing genetic variation*
361 Population clustering analyses demonstrate that regional CPB pest populations (western
362 US, eastern US, multiple lineages in the Midwestern US, and Europe) are genetically distinct,
363 with D-statistics suggesting limited ongoing gene flow. This alone supports previous
364 observations that insecticide resistance evolves locally among CPB populations [43, 51, 52, 55].
365 However, genome scans, using both outlier-based and environmental association-based methods,
366 provide clear signatures that adaptation occurs repeatedly across different populations of CPB.
367 Focusing on candidate genes of insecticide resistance, the considerable overlap among the two

368 methods suggests different genes are selected in different populations (see PC loadings in **Table**
369 **2**). Similarly, candidate insecticide resistance genes are constitutively upregulated, but in distinct
370 patterns in different populations (**Fig. 4**). Selected genes also encompass multiple resistance
371 mechanisms, including metabolic detoxification, target site resistance, and cuticular proteins,
372 which is broadly supported by gene set enrichment analyses across multiple tests (**Additional**
373 **File 1: Fig. S34**). These results are consistent with previous genetic studies that have
374 documented some of the same genes (see **Table 2; Additional File 2**) or mechanisms in CPB
375 resistance [43, 61-64]. We also identify compelling new candidates, such as the multiple ABC
376 transporters, as well as acetylcholine receptor β subunit (LDEC002850), which are likely linked
377 to neonicotinoid resistance phenotypes [65, 66]. Although our analysis focuses on pesticide
378 resistance candidates, we note that other interesting genes emerged from our study. In particular,
379 an octopamine receptor (LDEC006841), which was identified as a gene associated with pest
380 behavior in a comparative genomics analysis of CPB [67], is recovered as a significant target in
381 contrasts of Plains and pest populations in both *LFMM* and *PCAdapt*.

382 The observed patterns of repeated local adaptation are most consistent with polygenic
383 evolution from standing genetic variation. Using a more sophisticated haplotype-based method
384 (focusing on ~21% of the genome), we found 24 highly significant ($p < 0.0001$) haplotype
385 blocks suggestive of selective sweeps. Only one selected region exceeded 4 Kb in length (35
386 Kb), while the remaining regions averaged 3.1 Kb in length. These shorter sweep lengths are
387 similar to those found in at least one other insect pest, *Spodoptera frugiperda* [on average,
388 sweeps are 4.1 Kb: 68], but much smaller than the strong selective sweeps found in other
389 prominent cases of insecticide resistance [18, 69, 70]. Sweep lengths typically scale inversely
390 with the rate of recombination as a product of increasing effective population size, so we might

have expected sweep lengths similar to those found in insects with large population size (typically > 10 Kb in *D. melanogaster*) [71]. Instead, nearly all strong candidate selective events from *hapFLK* are more consistent with soft sweeps that recur in multiple geographically distant populations. Most importantly, the sweep regions in CPB occur on multiple branches of the population tree (on average, 8.16 branches). For soft sweeps to be a reasonable mechanism in the evolution of insecticide resistance, however, populations must maintain high levels of genomic variation. The host-shift to potato, coupled with agroecosystem invasion over a broad geographical scale, has been presumed to reduce genetic variation in CPB [47, 72]. However, despite having recently invaded agroecosystems, CPB pest populations exhibit only modest reductions in nucleotide diversity (though the loss of private alleles is more pronounced), and SMC++ reconstructions show recent population expansion. Dense population sampling of CPB at a landscape scale in the Midwestern U.S. also supports high levels of standing variation at a regional level [52].

Early population genetic research on insecticide resistance by J.F. Crow suggested that resistance evolution arises from polygenic standing variation [7]. Quantitative variation in insecticide tolerance, among populations and individuals within populations, is frequently observed, even under controlled laboratory conditions [73, 74]. Increasingly, genomics-based analyses are documenting how multiple genes contribute to quantitative resistance phenotypes [16, 75-77]. A combination of new mutations and recruitment of standing genetic variation probably occurs in many cases of adaptation, as it is evident that phenotypic traits are typically quantitative in nature (influenced by multiple loci), large-effect mutations often require compensatory fitness changes, and the most likely outcome of evolution in theoretical models is polygenic adaptation [78]. At the same time, the action of multiple genes and their regulatory

414 elements can be difficult to detect using population genomics methods, as it results in modest
415 changes in allele frequencies (under soft or partial selective sweeps) [79-81]. Further noise is
416 added because tests of selection are prone to false positives [82]. For example, *hapFLK* tends to
417 show a high false positive rate under population models with continuous migration or in those
418 that experience strong bottlenecks [83, 84]. However, while chance false positives might
419 contribute to our observed patterns, it is unlikely that so many would occur in candidate
420 insecticide resistance genes. Our results are in broad support with the emerging view that
421 polygenic architecture is common in insecticide resistance [10, 24, 85, 86]. Certainly of course,
422 expanded and more continuous population sampling of CPB will be needed to improve support
423 for this mode of selection and the complex genetic architecture underlying insecticide resistance.
424
425 *Selection on de novo mutation: is CPB mutation limited?*
426 Selection on novel mutations could lead to repeated evolution of resistance in pests with
427 intrinsically large population sizes (much greater than $>>10^6$) that are not considered mutation-
428 limited [29, 30]. Analyses of nucleotide diversity in protein-coding genes is known to range
429 widely in animal species and is most strongly correlated with reproductive strategy, with highly
430 fecund (r-selected) species like CPB having the greatest diversity [87]. Schoville *et al*. [58]
431 found evidence for a high rate of polymorphism in protein-coding regions (nucleotide diversity,
432 $\pi$, was ~0.01), suggesting a high level of standing genetic diversity relative to other insects.
433 Furthermore, CPB shows a higher rate of positive selection and greater levels of standing
434 variation compared to other species in the genus *Leptinotarsa* [88]. In this genome-wide
435 resequencing dataset, we identify an average nucleotide diversity of 0.005 within CPB,
436 irrespective of pest or non-pest status across different geographical regions. While this rate of

nucleotide diversity remains high relative to most vertebrates (median 0.0025), it is not exceptional among arthropods [median 0.0125 all sites, 0.00204 synonymous sites; 89]. Despite its super-pest status, CPB nucleotide diversity falls within the range of insect species (0.0023-0.0288). Among Coleoptera, CPB falls within the range of nucleotide diversity found in the bark beetles *Dendroctonus ponderosae* and *D. brevicornis* [0.0023 and 0.008, respectively; 90, 91] and is similar to the horned scarab beetle *Onthophagus taurus* [0.0056, 92]. Lepidopteran pest genomes are more polymorphic; five species of *Helicoverpa* range in nucleotide diversity from 0.004 to 0.010 in autosomal regions, with the super-pest *H. armigera* as the most polymorphic [93], while two sympatric stains of the pest species *Spodoptera frugiperda* range from 0.043 to 0.044 [94] and the closely related pest *S. litura* has nucleotide diversity as high as 0.016 [95].

These results suggest that CPB is not exceptional in terms of standing genetic diversity and the potential for rapid evolution. However, there is an upper limit on nucleotide diversity in species with large effective population sizes (e.g. *Drosophila melanogaster*), as nucleotide variation at neutral sites is removed as a result of selection on nearby linked sites [96]. In fact, patterns of reduced variation in species with large effective population sizes might reflect selection from recurrent adaptive mutation [97]. On the other hand, there is considerable debate about this interpretation, as reduced nucleotide diversity might alternatively arise from purifying selection acting on nearly neutral sites in the form of background selection [98, 99]. Distinguishing among these alternatives will require direct estimates of genome-wide mutation and recombination in CPB, in addition to improved sampling, as it is not yet clear that CPB is mutation-limited.

*Are key regulatory shifts contributing to resistance evolution*

Overexpression of multiple CYP, GST and ABC transporter genes is often associated with insecticide resistance and several studies have shown that trans-acting transcription factors may simultaneously regulate the expression of these targets [100-103]. Known trans-acting transcription factors involved in the xenobiotic detoxification pathway include *CncC*, *Maf-S*, *AhR*, *ARNT*, and *Met*. *CncC* forms a heterodimer with *Maf-S* to regulate multiple detoxification loci in many insects [104], including the highly resistant Long Island population of CPB [63] and deltamethrin resistant populations of the beetle *Tribolium castaneum* [105]. However, in comparing gene expression profiles of CPB populations throughout several growing regions, we found that expression levels of insecticide detoxification genes vary across populations, suggesting that varied transcriptional patterns are most-likely achieved through cis-regulatory evolution [106, 107]. In a similar analysis, comparison of transcriptomic profiles of *Anopheles gambiae* across Africa revealed the recruitment of many population-specific candidate insecticide resistance genes [108], suggesting cis-regulatory evolution of these pathways may be common among insect pests. Our data don't support the role of a single master-regulatory switch driving insecticide resistance, as we see no evidence for differential expression of key trans-acting transcription factors despite comparing insecticide resistant and susceptible populations. Furthermore, the role of different genes in metabolic resistance in CPB has been demonstrated by RNAi experiments where knockdown of different upregulated CYP genes restores susceptibility in different pesticide resistant CPB populations [109, 110]. Altogether, our results suggest that a simple upstream shift in *Cap-n-collar* expression is not sufficient to explain all cases of metabolic resistance in CPB and that, instead, additional cis-regulatory changes are required to account for the heterogeneity and diversity of resistance pattern among populations [111]. This is consistent with widespread evidence that cis-regulatory evolution is more common

in adaptive evolution, while trans-acting gene regulation is typically constrained by strong stabilizing selection [112, 113].

*Implications for pest management and novel control tactics*

Current resistance management strategies assume that the evolution of resistance is a rare event, caused by simple (single-gene) mutations [114], thereby ignoring the importance of alternative mechanisms involving multiple loci [115]. Resistance management models also assume that resistance is involved in fitness tradeoffs [116], a rationale underlying high dose/refuge strategies where gene flow from susceptible 'refuge' populations into insecticide-treated fields delays resistance evolution [117]. However, it is increasingly difficult to understand the rate of pesticide resistance evolution using conventional models of single, large-effect mutations in conferring a resistance phenotype [118]. Rates of insecticide resistance evolution in CPB are among the highest observed in agricultural pests [6] and often lead to failure unless implemented in an integrated pest management framework [119]. From our results, polygenic evolution from standing variation appears best explains this pattern, although we note alternative mechanisms were not investigated and could contribute to rapid evolutionary change. Notably, recent work in CPB has shown that changes in DNA methylation patterns might drive transgenerational epigenetic mechanisms of regulatory evolution that lead to pesticide resistance [120, 121].

As society faces the challenge of global food security, there is a prevailing view that insect pests have won the arms race involving conventional chemical pesticide control [3]. Novel chemical modes of action are needed to avoid target-site resistance, yet there is a high cost to such efforts, both in terms of development costs and environmental impacts. In addition, the

widespread emergence of cross-resistance, especially through metabolic detoxification, suggests novel modes of action may have limited efficacy and durability [122]. Although gene drives have been raised as promising novel control tactics [123, 124], most recent work in in CPB has focused on gene-targeted insecticides via the RNAi pathway [125]. Gene knock-down via RNAi could allow for highly effective, species-specific management if multiple genes are targeted simultaneously, and such products are currently under development [126]. How do genome-wide population genetic data shed light on RNAi implementation? Drawing on the mechanisms of evolution described in this paper, where standing variation is substantial, the likelihood of resistance evolution to RNAi might be high unless population-specific approaches are developed. One target site mutagenesis experiment in CPB has shown that mismatch rate of 3% or less still allows for effective gene target suppression [127], but clearly some CPB target genes would be problematic. Additionally, alternative pathways of RNAi resistance might emerge. For example, experiments with cell lines of CPB have shown that resistance evolving from mutations altering the uptake and transport of dsRNA [128]. CPB is known to utilize both *sid-1* transmembrane channel-mediated uptake and clathrin-mediated endocytosis in processing dsRNA [129], suggesting that there are multiple targets for resistance evolution in the RNAi pathway. Comparison of dsRNA efficacy among European CPB populations also suggested variation in the RNAi pathway itself (involving the multiple homologs of *dicer*, *argonaut*, and *staufen*) was more likely to evolve than at target site loci [130]. However, knockdown experiments in other Coleopteran pests have shown that loss of function of genes in the RNAi pathway impair development and reduce reproductive fitness [131], thus the inherent trade-off in resistance may be too great. One interesting RNAi resistance pathway involves selection on gut nuclease activity that alters the sensitivity of CPB to RNAi [132]. In other insects, such as Lepidopteran pests

529 [133], a single nuclease is responsible for dsRNA tolerance. Though clearly dsRNA provides a
530 novel mode of action for controlling CPB pests, the propensity to draw on reservoirs of standing
531 genetic variation to rapidly evolve suggests that multiple mechanisms of resistance are likely to
532 occur.
533
534 **Conclusions**
535 Understanding the molecular mechanisms underlying pesticide adaptation has become
536 increasingly important because of the widespread occurrence of the "pesticide treadmill"
537 phenomenon in agricultural pests [134], wherein the repeated and escalating use of pesticides,
538 and the search for new chemistries [119], is required to keep pace with pest evolution. We
539 provide clear evidence that polygenic resistance drawn from standing variation could explain
540 how insects rapidly overcome multiple classes of pesticides [135]. While the importance of
541 polygenic evolution from standing genetic variation remains mostly theoretical in the insecticide
542 resistance literature and has proven challenging to identify in empirical case studies [8],
543 polygenic resistance has also been broadly implicated in the evolution of herbicide resistance in
544 agricultural weeds [85] and antibiotic resistance in bacterial pathogens [136]. Here we provide
545 evidence that CPB evolves insecticide resistance repeatedly across agricultural regions, and often
546 at the same loci. An important future goal will be to understand how polygenic resistance
547 phenotypes spread among local pest populations, in order to refine integrated pest management
548 practices to maintain the efficacy and sustainability of novel control techniques.
549
550 **Methods**
551 ***Study Design and Aim***

552 We collected a geographically dispersed set of 88 samples was selected to maximize
553 information about genomic differentiation across the range of CPB (**Additional File 1: Fig. S1**
554 and **Table S1**). An additional 10 samples comprising nine species of *Leptinotarsa* were also
555 collected for relevant information on outgroup variation and possible sources of hybridization.
556 Within the 88 CPB samples, we sampled six geographically proximate pairs of resistant (R) and
557 susceptible (S) populations, 5 beetles per R/S site: Maine (R) and Vermont (S), New York (R)
558 and New Jersey (S), Maryland (R and S), Michigan (R and S), Wisconsin (R and S), Oregon (R
559 and S). The resistance status of beetles was ascertained by topical exposure to an insecticide
560 (imidacloprid) or from published records at those sites (see **Additional File 1** for detailed
561 methods). However, all pest populations have potentially evolved resistance to other insecticides,
562 as insecticide use was widespread starting in the late 1940s. As each diploid individual
563 represents N=2 genomes, our sample size exceeds the requirements for most population genomic
564 tests and allows for accurate estimation of frequencies for all but the most rare (and therefore,
565 presumably, less important) alleles in key potato growing regions of the United States [137].
566
567 ***Genomic Resequencing, Quality Control and Variant Calling***
568 High quality genomic DNA was isolated from adult beetle thoracic muscle tissue using
569 DNeasy Blood & Tissue kits (Qiagen) and then submitted to the University of Wisconsin-
570 Madison Biotechnology Center. Libraries were sequenced using paired-end, 125bp sequencing
571 on a HiSeq2500 sequencer (see **Additional File 1** for detailed methods). We predetermined
572 sequencing effort to yield >6x average coverage for each of our CPB genomes, a quantity
573 sufficient to identify SNPs with reasonable accuracy [138].

Each sample was demultiplexed prior to downstream analysis, and we followed GATK's "Best Practices" guidelines (https://software.broadinstitute.org/gatk/best-practices/). Using the *L. decemlineata* reference genome v1.0 [GCA_000500325.1; 58], we aligned demultiplexed reads using BWA v0.7.101 [139], and converted SAM files to BAM format using SAMTOOLS v1.3.12 [139]. We generated one uBAM file (i.e., unmapped BAM file) per forward-reverse pair of the fastq raw reads using FastqToSam and then marked Illumina adapters with MarkIlluminaAdapters, both functions available from PICARD v2.2.4 (https://github.com/broadinstitute/picard). We then reverted BAM files to fastq format with PICARD's SamToFastq, aligned the new fastq files to the reference genome with the BWA-mem algorithm and merged all alignments into one BAM file per sample with PICARD's MergeBamAlignment tool. We marked PCR and optical duplicates using PICARD's MarkDuplicates tool, but some of our samples were sequenced on multiple sequencer lanes. For these samples, we marked duplicates first at the lane level (i.e., per replicate), then at the sample level (merging duplicates into a unique BAM output). Finally, we realigned reads around insertions and deletions with GATK's RealignerTargetCreator and IndelRealigner tools. In order to assess the quality of our BAM files, we used GATK's Flagstat and DepthOfCoverage tools. Among our 88 CPB samples, three samples (two susceptible samples from Oregon: CPBWGS_59 and CPBWGS_63, and one susceptible sample from Vermont: CPBWGS_93) had few successfully mapped reads and were removed (**Additional File 1**: **Fig. S2** and **Table S1**).

Genotyping was split into two steps: per-individual variant calling, followed by joint genotyping. Variant calling was conducted with GATK's *HaplotypeCaller* tool, which generates a likelihood score for all reference sites (-ERC GVCF option), including non-variant sites. Two different joint genotyping procedures were performed: one excluding non-CPB samples ("CPB"

dataset; N=85), and one including non-CPB samples, but keeping only one susceptible and one resistant sample (chosen at random) for populations from Oregon, Wisconsin, Michigan, Maryland, New-Jersey/New-York and Vermont/Maine ("*Leptinotarsa*" dataset; N=50; **Additional File 1: Table S2)**. For joint genotyping, we employed variant quality score recalibration (VQSR) using a training dataset. VQSR is based on applying machine learning algorithms and clustering methods to examine the overlap of the raw call set and a training dataset. It is composed of two steps: 1. *ApplyRecalibration* describes the multi-dimensional annotation profile of variants and calculates (for each variant in both datasets) a new, well-calibrated quality score called VQSLOD (for "variant quality score log-odds"). 2. *ApplyRecalibration* uses VQSLOD to apply a new cutoff to retain only high likelihood variants from the call set, based on a proportion of the variants in the training set that are present in the call set (e.g. 99.9% to enhance sensitivity or 90% to enhance specificity). As this approach requires a well-validated, independent dataset to be used as a training set, we used 41,454 SNPs generated from a published genotyping-by-sequencing (GBS) experiment [52]. These data represent 188 samples from 24 Midwestern populations (**Additional File 1: Fig. S35)**, which were hard filtered for depth of coverage ≥10x, polymorphism in at least 30% of the individuals of each population, minor allele frequency ≥5%, and less than 20% missing genotypes across all individuals. We calculated VQSLODs based on the following annotations: QD, MQ, MQRankSum, ReadPosRankSum, FS, SOR, DP and InbreedingCoeff. The recalibrated score provides a continuous estimate for the probability of each variant, which can then be partitioned into quality tranches. Tranche plots for the "CPB" and "*Leptinotarsa*" datasets are based on a 90% threshold that maximizes specificity over sensitivity (**Additional File 1: Fig. S36**). Finally, we used GATK's *VariantsToTable* tool to assess the quality of our inferred SNP dataset. We

plotted the improvement in the distribution of QualByDepth (QD) following the VQSR procedure for the "CPB" dataset (**Additional File 1: Fig. S37**) using the *ggplot*2 package in R v3.6 [140, 141]. After removing sites representing mitochondrial DNA (1,756 total variant sites), our dataset contained 47,969,460 SNPs in the "CPB" dataset and 69,680,768 in the "*Leptinotarsa*" dataset. Some analyses, like demographic reconstruction, require analyses with neutrally evolving loci. To mitigate the effect of non-neutral loci in these analyses, we created an "intergenic CPB dataset" from our "CPB" dataset, by considering only SNPs located outside of known genes from the *L. decemlineata* Official Gene Set (OGS) v1.1 [58].

***Genomic Diversity***

To estimate the genetic diversity of populations, we used the "CPB" dataset (*i.e.* not including Mexican samples), removed multi-allelic SNPs and those SNPs with a MAF <5% (using GATK's *SelectVariants* tool; final SNP number = 17,599,906). For these analyses we consider paired populations (susceptible/resistant) and grouped individuals from the U.S. Plains region (Colorado, Nebraska, Kansas, Missouri, New Mexico and Texas). We estimated genome-wide nucleotide diversity ($\pi$) using a 10 Kb sliding window with VCFtools v0.1.15 [142]. We also estimated heterozygosity by calculating the inbreeding coefficient *F* for each individual, using the method of moments estimator in VCFtools. Individual estimates were then averaged per population, keeping susceptible and resistant individuals separate. Finally, we used the "*singletons"* function in VCFtools to calculate the number of singletons and private doubletons for each sample.

***Evolutionary Divergence***

In order to reconstruct the evolutionary origins of CPB populations, we conducted a phylogenomic analysis of the "*Leptinotarsa*" dataset. This dataset comprised 48 samples, including 10 *L. decemlineata* samples from Mexico, 28 *L. decemlineata* from the U.S. and Europe, and 10 samples of closely-related *Leptinotarsa* species. Since the dataset was quite large, we used only the first 100 scaffolds comprising 16,519,065 SNPs. We used SNPhylo v.20160204 [143] to construct a phylogeny, after pruning the SNPs for linkage disequilibrium (LD). In order to ensure the relative independence of the SNPs used in the analysis, we tested several values of the LD threshold parameter and selected 0.5 for downstream analyses, which resulted in 35,838 SNPs. The SNP data were concatenated and then aligned using MUSCLE v3.8.31 [144], and a maximum likelihood phylogeny was estimated using DNAML in PHYLIP v3.6 [145]. Support values for nodes in the tree were determined by bootstrap resampling 100 times.

To estimate population structure, we examined both the "CPB" dataset and the “intergenic CPB” dataset, but the results were biologically consistent. We used three approaches: classical $F_{ST}$ estimates between pairs of populations, principal components analysis using *PCAdapt* [146, 147] and ancestry analysis with *sNMF* (Frichot and François 2016). For the $F_{ST}$ analyses, we consider paired populations (susceptible/resistant) and Plains individuals (Colorado, Nebraska, Kansas, Missouri, New Mexico and Texas). All estimations of Weir and Cockerham's mean weighted fixation indices ($F_{ST}$) were done using 10 Kb windows in VCFtools [142]. For the PCAdapt analysis, we assessed the number of principal components using Cattell's rule [148]. We plotted the percentage of variance explained by the first 20 principal components in a screeplot and examined the point of inflection in the plot, above which additional terms provide diminishing returns in terms of explained variance. For *sNMF*, we inferred individual patterns of

ancestry by estimating ancestral population allele frequencies and admixture coefficients using the R package *LEA* [149]. We first converted our VCF files to PLINK's ped format using VCFtools, then to the geno file format using *LEA*'s *ped2geno* tool. We implemented 10 runs per k value and combined the different runs with CLUMPAK (http://clumpak.tau.ac.il/). We plotted cross-entropy values to assess the number of k values.

Finally, to test for possible admixture among ancestral populations during the invasion of CPB into agroecosystems, we assessed evidence of gene flow (i) from Mexican CPB populations and (ii) from other *Leptinotarsa* species, into the pest lineage. We used *Dsuite* [150] to calculate the genome-wide *D*-statistic (*D*) on allele frequency data, estimating the strength of introgression based on the ABBA/BABA test [151]. *D* ranges from zero (no introgression) to one (complete introgression), and is calculated using a set of three focal populations or taxa (P1, P2), P3) with one additional outgroup. For the first test (i), we only considered populations with sample sizes ≥ 5, grouping susceptible and resistant samples from Vermont/Maine, Maryland, Wisconsin and Oregon as $F_{ST}$ between these sub-populations was negligible (**Additional File 1: Table S5**), grouping samples from the "Plains" region (N=7), and grouping all CPB samples from Mexico (N = 10). We tested every possible focal trio of populations, retaining the lowest *D*-statistic for every given trio ($D_{min}$; a conservative estimate of *D*). We assessed whether *D* was significantly different from zero by calculating a *p*-value based on jackknifing. We analyzed the complete biallelic SNP dataset with the two *L. juncta* samples as outgroups (the results were almost identical when using *L. undecemlineata* instead; data not shown). For the second test (ii), we included the nine other *Leptinotarsa* species using the same dataset from the SNPhylo analysis, comprising the 100 first scaffolds and 16,519,065 SNPs. The dataset contained one susceptible and one resistant sample for each of the six paired populations. In order to analyze a balanced

dataset, we limited the "Plains" population to the samples from Colorado. We also created two Mexican populations, representing the two distinct Mexican clades recovered in our phylogeny: "Mexico_City" (containing two samples) and "Mexico_South" (containing the sample from Oaxaca and the one from Guerrero; **Additional File 1: Table S1**). We used *L. lineolata* as the outgroup, as it was recovered as the most basal and distantly-related taxon to the U.S. CPB clade.

***Demographic Analysis***

To reconstruct the population history of CPB, we used the CPB intergenic SNP dataset and employed two coalescent approaches: the stairway plot method [152] and SMC++ [153]. Demographic reconstructions rely on an accurate estimate of the mutation rate. Most estimates of nuclear mutation rate in insects fall into the range of $2 \times 10^{-9}$ to $7 \times 10^{-9}$ substitutions per site per generation [154, 155]. As there is no genome-wide mutation rate estimate for CPB or related beetles, we chose to use a mutation rate of $2.1 \times 10^{-9}$, estimated recently in the non-biting midge [156]. We set the generation time of 0.5/year (*i.e.* 2 generations per year) for all our samples. The Stairway plot approach relies on the calculation of the expected composite likelihood of a given site frequency spectrum (SFS), which reduces the computational burden of inferring population parameters. It is also suitable for estimating recent population histories with low coverage genomic data. We analyzed the resistant and susceptible paired populations, both separately and pooled together, and also considered a pooled sample from the Plains (Colorado, Nebraska, Kansas, Missouri, New Mexico and Texas), East (Florida, Tennessee, North Carolina, Virginia, Kentucky and Ohio) and Europe (Italy, Russia). For each population, we estimated the folded SFS in *dadi* [157], calculated for a genome length of 678Mbp [58], and used 200 bootstraps to assess confidence intervals. We then conducted the Stairway plot analysis with default

712 parameters and plotted the estimated median (and 95% confidence intervals) effective population
713 size (Ne) through time.

714 We chose SMC++ [153] as an alternative approach, as it incorporates estimates of
715 recombination and linkage disequilibrium (LD) in an SFS framework. Even though this method
716 is relatively computationally efficient, long run times prevented us from using the entire CPB
717 dataset. Due to the fragmented nature of the genome, we analyzed the longest 95 genomic
718 scaffolds of the CPB dataset, including all intergenic and non-intergenic biallelic SNPs. This
719 encompasses ~21% of the genome (~140MB out of ~670M bp, mean length of 1.9MB and all >
720 1MB) and contains 6,669,259 SNPs. This subset of highly contiguous reference genomic data
721 ensures we can accurately infer demography using the sequential Markov coalescent [158],
722 although we note that analyses considering all >10 Kb scaffolds (covering 95% of the genome)
723 yielded comparable results (i.e. same population sizes, same demographic events, same time-
724 scale; results not shown).

725

726 ***Selection Analyses***

727 We used three different approaches to study genomic signatures of selection: outlier
728 detection with *PCAdapt* [146, 147], genome-environment association with *LFMM* [159] and
729 haplotype-based tests using *hapFLK* [84]. To detect outlier SNPs using *PCAdapt*, Mahalanobis
730 distances were transformed into *p-values* and then the FDR was controlled by transforming the
731 *p-values* into *q*-values and considering an FDR of 0.01% ($\alpha$=0.0001). We initially filtered SNPs
732 using a minor allele frequency (MAF) of 0.05 and a conservative setting of K=10, but the
733 number of SNPs suggested a high rate of false positives (see **Additional File 1**). We therefore
734 refined our filtering steps using linkage disequilibrium clumping (choosing a window size of 500

735 SNPs and a squared-correlation coefficient threshold of 0.2). We examined the screeplot of the
736 principle components and, following Cattell's rule, selected K=6 as the optimal clustering level.
737 Finally, we adjusted the dataset by setting a more conservative MAF setting of 0.1.
738 As an alternative genome scan approach, we employed a genome-environmental
739 association method using latent factor mixed models [LFMM; 159]. This test identifies SNP
740 allele frequencies that are significantly associated with environmental predictor variables, while
741 simultaneously modeling the confounding effect of population structure as latent factors. To
742 account for the population structure observed in our data, we modeled k = 6 latent factors, as
743 suggested by our *PCAdapt* and *sNMF* results. We adjusted *p*-values by an empirically-
744 determined genomic inflation factor, while controlling the false discovery rate at 0.01%. We
745 explored five different environmental variables: elevation, precipitation, minimum temperature
746 in the coldest month and potato land cover. We reasoned that genes containing SNPs associated
747 with climate variables could be related to *L. decemlineata* adaptation to northern climates during
748 range expansion, while associations with potato land cover might reveal genes responding to
749 selective pressures faced in potato agroecosystems, such as novel host plants, natural enemy
750 pressure and insecticide exposure. We obtained historic, county-level potato land cover data
751 (between 1850 and 2012), as detailed in Crossley *et al*. [160]. For latitude, elevation,
752 precipitation and minimum temperature in the coldest month, we obtained the data from the
753 PRISM climate group (http://prism.oregonstate.edu). For each environmental variable, we took
754 the average value within a 75 km radius around each sampling site, using functions available in
755 the *rgdal* and *raster* packages in R [161, 162]. For potato land cover, we summarized the average
756 proportion of area planted with potato within a 75 km radius of each sample site [163].

Finally, to integrate information from linked SNPs in tests of selection, we used the haplotype frequency-based method *hapFLK* [84]. This method has been shown to be relatively robust to confounding effects of population structure and variable population size. It also allows selection events to be pinpointed to specific branches of the population tree. For each identified signature of selection, a local tree is re-estimated using significant SNPs, constrained by the overall topology of the population tree. Statistical significance is computed for the difference between the branch lengths estimated from the focal region and from the global tree. We analyzed the first (longest) 95 genomic scaffolds of the CPB dataset (representing ~21% of the genome), including all biallelic SNPs. From a VCF file produced with GATK's *SelectVariants* function, we produced a ped file and associated map file with VCFtools' --*plink* function. In order to use the multi-point linkage disequilibrium model, *hapFLK* needs the number of haplotype clusters (K) to be specified and a population tree. We compared *hapFLK* results for different K values, ultimately selecting K=20 to minimize imputation errors (see **Additional File 1** for detailed methods; **Additional File 1: Fig. S38**). The population tree was estimated from a kinship matrix (**Additional File 1**: **Fig. S39**). We standardized *hapFLK* values and computed corresponding *p-values* from a standard normal distribution, examining results at three nominal levels ($\alpha$ = 0.01, 0.001 and 0.0001; **Additional File 1**: **Fig. S40**). We grouped significant SNPs into selected regions, where each region was separated by at least 1 Kb up- and downstream.

***Candidate Genes and Gene Network Analysis***

For each selection test, we obtained functional information for candidate SNPs using manual annotation of the OGS supplemented by Blast2GO annotations [164], which have been previously published [52, 58]. To develop a list of candidate insecticide resistance genes,

Crossley *et al.* [52] identified 664 genes associated with processes or functions potentially linked to known mechanisms of insecticide resistance (**Additional File 1: Table S14**): metabolic detoxification including cytochrome p450s (CYPs), esterases, Glutathione S-transferases (GSTs) and ATP-binding cassette (ABC) transporters [165]; target-site insensitivity including most of the modes of actions classified by the *Insecticide Resistance Action Committee* (IRAC; http://www.irac-online.org/modes-of-action/) such as TRPC channels, sodium and calcium channels, glutamate receptors, acetylcholine receptors, etc.; and reduced cuticular penetration, including genes involved in chitin production or cuticle development.

One frequent concern in analyses of large datasets is the inclusion of false positives (Type I error) resulting from the large number of statistical tests. Frequently, this is resolved by adjusting *p*-values to more conservative values, for example by implementing multiple testing recalibrations such as Bonferroni or the Benjamini-Hochberg false discovery rate. However, in functional genomic studies, gene lists provide objective hypotheses that can be easily assessed in follow-up studies, while removing them based on statistical criteria can sometimes be challenging [166]. Although we employ stringent statistical criteria in all tests, we additionally leverage information by comparing gene lists from different tests. As we expect that both regulatory and structural changes might lead to genetic adaptation, we tested for over-representation of specific gene networks in selection tests, gene expression analyses, and combinations of both approaches. Given multiple data sources, overlap in gene identity and function provides a measure of support for repeated evolution and polygenic adaptation. We curated gene ontology terms associated with significant genes in lists from genome-wide selection tests and differential expression tests, and used a one-sided hypergeometric Fisher's Exact test [167, 168] to test for over-representation (enrichment) of gene ontology terms, with *p*-

803 value < 0.05 used as the statistical significance threshold. To further refine this analysis, we used
804 REVIGO [169], a clustering algorithm that relies on semantic similarity measures, to summarize
805 the list of enriched gene ontology terms. Gene ontology terms associated with biological
806 processes, cellular components, and molecular function were separately clustered using the
807 simRel score for functional similarity, allowing for redundancy in similar terms up to a value of
808 0.7 before removal, and then compared to the UniProt database to find the percentage of genes
809 annotated with each gene ontology term. The results were visualized using a CirGO plot [170].
810 To provide a context for interpreting these results, we used our list of candidate genes
811 (**Additional File 1: Table S14**) to generate a CirGO plot of gene ontology terms associated with
812 insecticide resistance (**Additional File 1: Fig. S41**). Major biological processes include response
813 to insecticide/response to oxidative stress, endocytosis, glycerolipid metabolism, sensory
814 perception, and DNA integration. Major cellular components include the plasma
815 membrane/integral component of the membrane and transcription factor complex. Finally, major
816 molecular functions include metallocarboxypeptidase activity, monooxygenase activity,
817 acetylcholine binding, lipid binding, DNA polymerase binding, tetracycline transporter activity,
818 chromatin binding, chitin binding, and structural component of cuticle.
819
820 ***Gene Expression Analyses***
821 In order to test for regulatory evolution, we compared gene expression data from RNA
822 sequencing (RNAseq) experiments across the geographical range of CPB, including original data
823 from the Plains region (a Colorado population) and previously published pest CPB population
824 samples (see **Additional File 1: Table S11**). The Colorado population was raised under
825 greenhouse conditions on potato plants (~25$^{o}$C, 16:8 light:dark cycle), but represents the first

826 generation derived from wild-caught adults feeding on *Solanum rostratum*. RNAseq studies are
827 recognized as robust estimators of whole-genome gene expression profiles [171] that are highly
828 responsive to experimental conditions [172]. As the original experiments varied in their design,
829 we conducted a set of initial analyses to determine if sampling issues might bias geographical
830 comparisons. First, we assessed whether sampling of an overwintering (post-diapause)
831 population versus a summer (non-diapausing) generation at the same field in Wisconsin altered
832 gene expression patterns ($1^{st}$ generation versus $2^{nd}$ generation). Second, we examined whether
833 direct exposure to imidacloprid was necessary to induce insecticide resistance gene expression
834 responses, by comparing a set of lab-reared individuals from Wisconsin, Oregon and Long Island
835 populations (control versus an imidacloprid-induction treatment). Third, we compared samples
836 of larvae and adults from a susceptible population in Wisconsin. Based on these comparisons, we
837 determined that regional population differences could be compared for adults from field
838 collected populations irrespective of generation sampled, but lab reared larvae needed to be
839 compared separately. We compared constitutive levels of gene expression in six adult
840 populations: Colorado (CO), Wisconsin (WI), Michigan (MI), New York (NY), New Jersey
841 (NJ), and eastern Canada (CAN). We compared constitutive levels of gene expression in four
842 larval populations: Oregon (OR), Wisconsin (WI), New York (NY), and New Jersey (NJ). As
843 some of the adult samples were sequenced as pair-end and single-end reads, we analyzed only
844 the first read of a set of pair-end samples representing CO and WI.
845 We aligned short read data from each sample to the *L. decemlineata* reference genome
846 using HISAT2 [173]. SAMTOOLS was used to convert *sam* files to *bam* files. Read counts per
847 gene per sample were generated using the function *featureCounts* available in the *Rsubread*
848 package [174], with reference to the *L. decemlineata* OGS. Using the resulting counts, we

evaluated evidence for differential gene expression for each region using DESEQ2 [175]. DESEQ2 first estimates the dispersion among a set of replicated samples and then the logarithmic fold change of transcript counts among sample groups. It then employs a generalized linear model based on the negative binomial distribution of transcript counts and a binomial Wald statistic to test for differences among experimental contrasts. We first trimmed the read count matrix to remove genes with less than five reads and then conducted the differential expression analysis. We retained differentially expressed genes if read counts were > 2-fold and the significance level $\alpha = 5\%$ was reached. We then adjusted the false discovery rate to 1% level, using a Benjamini-Hochberg correction [176]. Heatmaps of differentially expressed genes were generated in R using the *pheatmap* package [177]. As described previously, we curated gene ontology terms associated with significant genes and used a one-sided hypergeometric Fisher's Exact test to test for over-representation (enrichment) of gene ontology terms, with a *p*-value < 0.05 used as the statistical significance threshold.

**Declarations**

The authors wish to thank the reviewers and editorial staff for their assistance with our manuscript. No ethics approval was required for this research and the authors declare no competing interests. All genomic data have been made publicly available at NCBI (Bioproject PRJNA580490). Funding was provided by a USDA NIFA AFRI Exploratory Grant (2015-67030-23495), a USDA-National Potato Council award (58-5090-7-073) and two Hatch Awards (WIS02004 and VT-H02010), in addition to support from Wisconsin Potato and Vegetable Growers Association. The authors thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing facilities and services. Finally, we thank Margarethe

872 Brummerman for help in sampling *Leptinotarsa* species, and an extensive network of collectors
873 that helped us to collect beetles in the U.S, Mexico, and Europe.

874

875 **Supplementary Information**

876 **Additional File 1** (.pdf): Supplementary Methods and Results. Contains detailed methodology
877 and results from analytical methods, including Supplemental Figures S1-S44 and Supplemental
878 Tables S1-S15.

879 **Additional File 2** (.xlsx): Significant genes from selection tests and gene expression tests.

880

881 **References**

882 1. Sharma S, Kooner R, Arora R: **Insect pests and crop losses.** In *Breeding Insect Resistant*
883 *Crops for Sustainable Agriculture.* Edited by Arora R, Sandhu S. Singapore: Springer; 2017:
884 45-66
885 2. Chen YH, Schoville SD: **Editorial overview: Ecology: Ecological adaptation in**
886 **agroecosystems: novel opportunities to integrate evolutionary biology and agricultural**
887 **entomology.** *Current Opinion in Insect Science* 2018, **26:**iv-viii.
888 3. Gould F, Brown ZS, Kuzma J: **Wicked evolution: Can we address the sociobiological**
889 **dilemma of pesticide resistance?** *Science* 2018, **360:**728-732.
890 4. Lewis WJ, Van Lenteren J, Phatak SC, Tumlinson J: **A total system approach to**
891 **sustainable pest management.** *Proceedings of the National Academy of Sciences* 1997,
892 **94:**12243-12248.

5. McCaffery A, Nauen R: **The insecticide resistance action committee (IRAC): public responsibility and enlightened industrial self-interest.** *Outlooks on Pest Management* 2006, **17:**11-14.

6. Brevik K, Schoville SD, Mota-Sanchez D, Chen YH: **Pesticide durability and the evolution of resistance: A novel application of survival analysis.** *Pest management science* 2018.

7. Crow JF: **Genetics of insect resistance to chemicals.** *Annual review of entomology* 1957, **2:**227-246.

8. Pélissié B, Crossley MS, Cohen Z, Schoville SD: **Rapid evolution in insect pests: the importance of space and time in population genomics studies.** *Current opinion in insect science* 2018.

9. ffrench-Constant RH, Bass C: **Does resistance really carry a fitness cost?** *Current opinion in insect science* 2017, **21:**39-46.

10. Ffrench-Constant RH: **The molecular genetics of insecticide resistance.** *Genetics* 2013, **194:**807.

11. Labbé P, Berticat C, Berthomieu A, Unal S, Bernard C, Weill M, Lenormand T: **Forty years of erratic insecticide resistance evolution in the mosquito Culex pipiens.** *PLoS genetics* 2007, **3:**e205.

12. Kamdem C, Fouet C, Gamez S, White BJ: **Pollutants and insecticides drive local adaptation in African malaria mosquitoes.** *Molecular biology and evolution* 2017, **34:**1261-1275.

914 13. Daborn P, Yen J, Bogwitz M, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D,
915 Batterham P: **A single P450 allele associated with insecticide resistance in Drosophila.**
916 *Science* 2002, **297:**2253-2256.
917 14. Tabashnik BE, Brevault T, Carriere Y: **Insect resistance to Bt crops: lessons from the first**
918 **billion acres.** *Nature Biotechnology* 2013, **31:**510-521.
919 15. Hoy MA: **Myths, models and mitigation of resistance to pesticides.** *Philosophical*
920 *Transactions of the Royal Society of London B: Biological Sciences* 1998, **353:**1787-1795.
921 16. Cheng T, Wu J, Wu Y, Chilukuri RV, Huang L, Yamamoto K, Feng L, Li W, Chen Z, Guo
922 H: **Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid**
923 **pest.** *Nature ecology & evolution* 2017, **1:**1747.
924 17. Anderson CJ, Oakeshott JG, Tay WT, Gordon KH, Zwick A, Walsh TK: **Hybridization and**
925 **gene flow in the mega-pest lineage of moth, *Helicoverpa*.** *Proceedings of the National*
926 *Academy of Sciences* 2018, **115:**5034-5039.
927 18. Calla B, Demkovich M, Siegel JP, Viana JPG, Walden KKO, Robertson HM, Berenbaum
928 MR: **Selective sweeps in a nutshell: the genomic footprint of rapid insecticide resistance**
929 **evolution in the almond agroecosystem.** *Genome Biology and Evolution* 2020.
930 19. Lucas ER, Miles A, Harding NJ, Clarkson CS, Lawniczak MK, Kwiatkowski DP, Weetman
931 D, Donnelly MJ, Consortium AgG: **Whole-genome sequencing reveals high complexity of**
932 **copy number variation at insecticide resistance loci in malaria mosquitoes.** *Genome*
933 *research* 2019, **29:**1250-1261.
934 20. Fritz ML, DeYonke AM, Papanicolaou A, Micinski S, Westbrook J, Gould F:
935 **Contemporary evolution of a Lepidopteran species, Heliothis virescens, in response to**
936 **modern agricultural practices.** *Molecular ecology* 2018, **27:**167-181.

21. Guillemaud T, Ciosi M, Lombaert E, Estoup A: **Biological invasions in agricultural settings: insights from evolutionary biology and population genetics.** *Comptes rendus biologies* 2011, **334:**237-246.

22. Kirk H, Dorn S, Mazzi D: **Molecular genetics and genomics generate new insights into invertebrate pest invasions.** *Evolutionary Applications* 2013, **6:**842-856.

23. Lee CE: **Evolutionary genetics of invasive species.** *Trends in ecology & evolution* 2002, **17:**386-391.

24. Hawkins NJ, Bass C, Dixon A, Neve P: **The evolutionary origins of pesticide resistance.** *Biological Reviews* 2019, **94:**135-155.

25. Georghiou GP, Taylor CE: **Factors influencing the evolution of resistance.** *Pesticide resistance: strategies and tactics for management* 1986**:**157-169.

26. May RM, Dobson AP: **Population dynamics and the rate of evolution of pesticide resistance.** *Pesticide resistance: strategies and tactics for management* 1986**:**170-193.

27. Hairston NG, Ellner SP, Geber MA, Yoshida T, Fox JA: **Rapid evolution and the convergence of ecological and evolutionary time.** *Ecology Letters* 2005, **8:**1114-1127.

28. Whitehead A, Clark BW, Reid NM, Hahn ME, Nacci D: **When evolution is the solution to pollution: Key principles, and lessons from rapid repeated adaptation of killifish (Fundulus heteroclitus) populations.** *Evolutionary applications* 2017, **10:**762-783.

29. Cutter AD, Jovelin R, Dey A: **Molecular hyperdiversity and evolution in very large populations.** *Molecular ecology* 2013, **22:**2074-2095.

30. Barton N: **Understanding adaptation in large populations.** *PLoS genetics* 2010, **6:**e1000987.

959 31. López-Maury L, Marguerat S, Bähler J: **Tuning gene expression to changing**
960 **environments: from rapid responses to evolutionary adaptation.** *Nature Reviews*
961 *Genetics* 2008, **9:**583.
962 32. Roelofs D, Morgan J, Stürzenbaum S: **The significance of genome-wide transcriptional**
963 **regulation in the evolution of stress tolerance.** *Evolutionary Ecology* 2010, **24:**527-539.
964 33. Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, Karchner SI,
965 Hahn ME, Nacci D, Oleksiak MF: **The genomic landscape of rapid repeated evolutionary**
966 **adaptation to toxic pollution in wild fish.** *Science* 2016, **354:**1305-1308.
967 34. Dermauw W, Wybouw N, Rombauts S, Menten B, Vontas J, Grbić M, Clark RM, Feyereisen
968 R, Van Leeuwen T: **A link between host plant adaptation and pesticide resistance in the**
969 **polyphagous spider mite *Tetranychus urticae*.** *Proceedings of the National Academy of*
970 *Sciences* 2013, **110:**E113–E122.
971 35. Messer PW, Petrov DA: **Population genomics of rapid adaptation by soft selective**
972 **sweeps.** *Trends in ecology & evolution* 2013, **28:**659-669.
973 36. Pujol B, Blanchet S, Charmantier A, Danchin E, Facon B, Marrot P, Roux F, Scotti I,
974 Teplitsky C, Thomson CE: **The missing response to selection in the wild.** *Trends in*
975 *ecology & evolution* 2018, **33:**337-346.
976 37. Barton NH, Keightley PD: **Understanding quantitative genetic variation.** *Nature Reviews*
977 *Genetics* 2002, **3:**11-21.
978 38. De Carvalho D, Ingvarsson PK, Joseph J, Suter L, Sedivy C, Macaya-Sanz D, Cottrell J,
979 Heinze B, Schanzer I, Lexer C: **Admixture facilitates adaptation from standing variation**
980 **in the European aspen (*Populus tremula* L.), a widespread forest tree.** *Molecular Ecology*
981 2010, **19:**1638-1650.

982 39. Corrêa AS, Cordeiro EM, Omoto C: **Agricultural insect hybridization and implications**
983 **for pest management.** *Pest management science* 2019.
984 40. Hardy NB, Peterson DA, Ross L, Rosenheim JA: **Does a plant-eating insect's diet govern**
985 **the evolution of insecticide resistance? Comparative tests of the pre-adaptation**
986 **hypothesis.** *Evolutionary applications* 2018, **11:**739-747.
987 41. Rane RV, Ghodke AB, Hoffmann AA, Edwards OR, Walsh TK, Oakeshott JG: **Detoxifying**
988 **enzyme complements and host use phenotypes in 160 insect species.** *Current opinion in*
989 *insect science* 2019.
990 42. Láruson ÁJ, Yeaman S, Lotterhos KE: **The importance of genetic redundancy in**
991 **evolution.** *Trends in Ecology & Evolution* 2020, **35:**809-822.
992 43. Alyokhin A, Baker M, Mota-Sanchez D, Dively G, Grafius E: **Colorado potato beetle**
993 **resistance to insecticides.** *American Journal of Potato Research* 2008, **85:**395-413.
994 44. Izzo VM, Hawthorne DJ, Chen YH: **Geographic variation in winter hardiness of a**
995 **common agricultural pest, *Leptinotarsa decemlineata*, the Colorado potato beetle.**
996 *Evolutionary Ecology* 2014, **28:**505-520.
997 45. Petek M, Turnšek N, Gašparič MB, Novak MP, Gruden K, Slapar N, Popovič T, Štrukelj B,
998 Jongsma MA: **A complex of genes involved in adaptation of *Leptinotarsa decemlineata***
999 **larvae to induced potato defense.** *Archives of Insect Biochemistry and Physiology* 2012,
1000 **79:**153-181.
1001 46. Izzo VM, Chen YH, Schoville SD, Wang C, Hawthorne DJ: **Origin of pest lineages of the**
1002 **Colorado potato beetle (Coleoptera: Chrysomelidae).** *Journal of economic entomology*
1003 2018, **111:**868-878.
1004 47. Lu W, Lazell J: **The voyage of the beetle.** *Natural history* 1996, **105:**36-39.

48. Walsh BD: **The new potato bug.** *Practical Entomologist* 1866, **2:**13-16.

49. Grapputo A, Boman S, Lindstroem L, Lyytinen A, Mappes J: **The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations.** *Molecular Ecology* 2005, **14:**4207-4219.

50. Izzo V, Chen YH, Schoville SD, Wang C, Hawthorne DJ: **Origin of pest lineages of the Colorado potato beetle, *Leptinotarsa decemlineata*.** *bioRxiv* 2017**:**156612.

51. Szendrei Z, Grafius E, Byrne A, Ziegler A: **Resistance to neonicotinoid insecticides in field populations of the Colorado potato beetle (Coleoptera: Chrysomelidae).** *Pest Management Science* 2012, **68:**941-946.

52. Crossley MS, Chen YH, Groves RL, Schoville SD: **Landscape genomics of Colorado potato beetle provides evidence of polygenic adaptation to insecticides.** *Molecular Ecology* 2017, **26:**6284-6300.

53. Hitchner EM, Kuhar TP, Dively GP, Youngman RR, Philips CR, Anderson TD: **Baseline toxicity and field efficacy of metaflumizone on Colorado potato beetle (Coleoptera: Chrysomelidae).** *Journal of economic entomology* 2012, **105:**207-213.

54. Rondon SI: **Pest management strategies for potato insect pests in the Pacific Northwest of the United States.** In *Insecticides - Pest Engineering.* Edited by Perveen F: InTech 2012: 309-333

55. Dively GP, Crossley MS, Schoville SD, Steinhauer N, Hawthorne DJ: **Regional differences in gene regulation may underlie patterns of sensitivity to novel insecticides in *Leptinotarsa decemlineata*.** *Pest Management Science* 2020, **76:**4278-4285.

56. Hsiao TH: **Host plant adaptations among geographic populations of the Colorado potato beetle.** *Entomologia experimentalis et applicata* 1978, **24:**437-447.

57. Hsiao TH: **Host specificity, seasonality and bionomics of *Leptinotarsa* beetles.** In *Biology of chrysomelidae.* Springer; 1988: 581-599

58. Schoville SD, Chen YH, Andersson MN, Benoit JB, Bhandari A, Bowsher JH, Brevik K, Cappelle K, Chen M-JM, Childers AK, et al: **A model species for agricultural pest genomics: the genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae).** *Scientific Reports* 2018, **8:**1931.

59. Casagrande R: **The Colorado potato beetle: 125 years of mismanagement.** *Bulletin of the Entomological Society of America* 1987, **33:**142-150.

60. Edgerton J: **Potato insects.** *Prairie Farmer* 1861, **8:**116.

61. Hawthorne DJ: **Quantitative trait locus mapping of pyrethroid resistance in Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae).** *Journal of economic entomology* 2003, **96:**1021-1030.

62. Clements J, Schoville S, Peterson N, Lan Q, Groves RL: **Characterizing molecular mechanisms of imidacloprid resistance in select populations of *Leptinotarsa decemlineata* in the Central Sands region of Wisconsin.** *PLoS ONE* 2016, **11:**e0147844.

63. Gaddelapati SC, Kalsi M, Roy A, Palli SR: **Cap'n'collar C regulates genes responsible for imidacloprid resistance in the Colorado potato beetle, *Leptinotarsa decemlineata*.** *Insect Biochemistry and Molecular Biology* 2018, **99:**54-62.

64. Rinkevich FD, Su C, Lazo TA, Hawthorne DJ, Tingey WM, Naimov S, Scott JG: **Multiple evolutionary origins of knockdown resistance (kdr) in pyrethroid-resistant Colorado potato beetle, *Leptinotarsa decemlineata*.** *Pesticide Biochemistry and Physiology* 2012, **104:**192-200.

65. Tan J, Salgado VL, Hollingworth RM: **Neural actions of imidacloprid and their involvement in resistance in the Colorado potato beetle, *Leptinotarsa decemlineata* (Say).** *Pest Management Science: formerly Pesticide Science* 2008, **64:**37-47.

66. Qu Y, Chen J, Li C, Wang Q, Guo W, Han Z, Jiang W: **The subunit gene Ldα1 of nicotinic acetylcholine receptors plays important roles in the toxicity of imidacloprid and thiamethoxam against *Leptinotarsa decemlineata*.** *Pesticide biochemistry and physiology* 2016, **127:**51-58.

67. Cohen ZP, Brevik K, Chen YH, Hawthorne DJ, Weibel BD, Schoville SD: **Elevated rates of positive selection drive the evolution of pestiferousness in the Colorado potato beetle (*Leptinotarsa decemlineata*, Say).** *Molecular Ecology* 2021, **30:**237-254.

68. Nam K, Nhim S, Robin S, Bretaudeau A, Nègre N, d'Alençon E: **Positive selection alone is sufficient for whole genome differentiation at the early stage of speciation process in the fall armyworm.** *BMC Evolutionary Biology* 2020, **20:**152.

69. Weedall GD, Riveron JM, Hearn J, Irving H, Kamdem C, Fouet C, White BJ, Wondji CS: **An Africa-wide genomic evolution of insecticide resistance in the malaria vector *Anopheles funestus* involves selective sweeps, copy number variations, gene conversion and transposons.** *PLOS Genetics* 2020, **16:**e1008822.

70. Schlenke TA, Begun DJ: **Strong selective sweep associated with a transposon insertion in *Drosophila simulans*.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101:**1626-1631.

71. Garud NR, Messer PW, Petrov DA: **Detection of hard and soft selective sweeps from *Drosophila melanogaster* population genomic data.** *bioRxiv* 2020.

72. Zehnder G, Sandall L, Tisler A, Powers T: **Mitochondrial DNA diversity among 17 geographic populations of *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae).** *Annals of the Entomological Society of America* 1992, **85:**234-240.

73. Tabashnik BE, Cushing NL: **Quantitative genetic analysis of insecticide resistance: variation in fenvalerate tolerance in a diamondback moth (Lepidoptera: Plutellidae) population.** *Journal of economic entomology* 1989, **82:**5-10.

74. Kobiela ME, Snell-Rood EC: **Genetic variation influences tolerance to a neonicotinoid insecticide in 3 butterfly species.** *Environmental Toxicology and Chemistry* 2020, **39:**2228-2236.

75. Faucon F, Dusfour I, Gaude T, Navratil V, Boyer F, Chandre F, Sirisopa P, Thanispong K, Juntarajumnong W, Poupardin R: **Unravelling genomic changes associated with insecticide resistance in the dengue mosquito Aedes aegypti by deep targeted sequencing.** *Genome research* 2015.

76. Dermauw W, Wybouw N, Rombauts S, Menten B, Vontas J, Grbić M, Clark RM, Feyereisen R, Van Leeuwen T: **A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae*.** *Proceedings of the National Academy of Sciences* 2013, **110:**E113-E122.

77. Barghi N, Hermisson J, Schlötterer C: **Polygenic adaptation: a unifying framework to understand positive selection.** *Nature Reviews Genetics* 2020, **21:**769-781.

78. Wellenreuther M, Hansson B: **Detecting polygenic evolution: problems, pitfalls, and promises.** *Trends in Genetics* 2016, **32:**155-164.

79. Messer PW, Ellner SP, Hairston NG: **Can population genetics adapt to rapid evolution?** *Trends in Genetics* 2016, **32:**408-418.

80. Stephan W: **Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation.** *Molecular ecology* 2016, **25:**79-88.

81. Teshima KM, Coop G, Przeworski M: **How reliable are empirical genomic scans for selective sweeps?** *Genome Research* 2006, **16:**702-712.

82. Mallick S, Gnerre S, Muller P, Reich D: **The difficulty of avoiding false positives in genome scans for natural selection.** *Genome research* 2009, **19:**922-933.

83. Vatsiou AI, Bazin E, Gaggiotti OE: **Detection of selective sweeps in structured populations: a comparison of recent methods.** *Molecular ecology* 2016, **25:**89-103.

84. Fariello MI, Boitard S, Naya H, SanCristobal M, Servin B: **Detecting signatures of selection through haplotype differentiation among hierarchically structured populations.** *Genetics* 2013, **193:**929-941.

85. Kreiner JM, Stinchcombe JR, Wright SI: **Population genomics of herbicide resistance: adaptation via evolutionary rescue.** *Annual Review of Plant Biology* 2018, **69:**611-635.

86. Schmidt JM, Battlay P, Gledhill-Smith RS, Good RT, Lumb C, Fournier-Level A, Robin C: **Insights into DDT resistance from the *Drosophila melanogaster* genetic reference panel.** *Genetics* 2017, **207:**1181-1193.

87. Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, Faivre N: **Comparative population genomics in animals uncovers the determinants of genetic diversity.** *Nature* 2014, **515:**261.

88. Cohen ZP, Brevik K, Chen YH, Hawthorne DJ, Weibel BD, Schoville SD: **Elevated rates of positive selection drive the evolution of pestiferousness in the Colorado potato beetle (*Leptinotarsa decemlineata*, Say).** *bioRxiv* 2019**:**870543.

89. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, Venkat A, Andolfatto P, Przeworski M: **Revisiting an old riddle: what determines genetic diversity levels within species?** *PLOS Biology* 2012, **10:**e1001388.

90. Bracewell R, Vanderpool D, Good J, Six D: **Cascading speciation among mutualists and antagonists in a tree–beetle–fungi interaction.** *Proceedings of the Royal Society B: Biological Sciences* 2018, **285:**20180694.

91. Keeling CI, Yuen MM, Liao NY, Docking TR, Chan SK, Taylor GA, Palmquist DL, Jackman SD, Nguyen A, Li M: **Draft genome of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major forest pest.** *Genome biology* 2013, **14:**R27.

92. Choi JH, Kijimoto T, Snell-Rood E, Tae H, Yang Y, Moczek AP, Andrews J: **Gene discovery in the horned beetle Onthophagus taurus.** *BMC Genomics* 2010, **11**.

93. Anderson CJ, Oakeshott JG, Tay WT, Gordon KHJ, Zwick A, Walsh TK: **Hybridization and gene flow in the mega-pest lineage of moth, Helicoverpa.** *Proceedings of the National Academy of Sciences* 2018, **115:**5034-5039.

94. Kiwoong N, Nhim S, Robin S, Bretaudeau A, Negre N: **Genomic differentiation is initiated without physical linkage among targets of divergent selection in Fall armyworms.** *bioRxiv* 2018**:**452870.

95. Cheng T, Wu J, Wu Y, Chilukuri RV, Huang L, Yamamoto K, Feng L, Li W, Chen Z, Guo H, et al: **Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest.** *Nature Ecology & Evolution* 2017, **1:**1747-1756.

96. Corbett-Detig RB, Hartl DL, Sackton TB: **Natural selection constrains neutral diversity across a wide range of species.** *PLOS Biology* 2015, **13:**e1002112.

1140 97. Karasov T, Messer PW, Petrov DA: **Evidence that adaptation in *Drosophila* is not limited**
1141 **by mutation at single sites.** *PLoS Genet* 2010, **6:**e1000924.
1142 98. Galtier N: **Adaptive protein evolution in animals and the effective population size**
1143 **hypothesis.** *PLOS Genetics* 2016, **12:**e1005774.
1144 99. Campos JL, Zhao L, Charlesworth B: **Estimating the parameters of background selection**
1145 **and selective sweeps in Drosophila in the presence of gene conversion.**
1146 *Proceedings of the National Academy of Sciences* 2017, **114:**E4762-E4771.
1147 100. Peng T, Chen X, Pan Y, Zheng Z, Wei X, Xi J, Zhang J, Gao X, Shang Q: **Transcription**
1148 **factor aryl hydrocarbon receptor/aryl hydrocarbon receptor nuclear translocator is**
1149 **involved in regulation of the xenobiotic tolerance-related cytochrome P450 CYP6DA2**
1150 **in Aphis gossypii Glover.** *Insect molecular biology* 2017, **26:**485-495.
1151 101. Smith LB, Tyagi R, Kasai S, Scott JG: **CYP-mediated permethrin resistance in *Aedes***
1152 ***aegypti* and evidence for trans-regulation.** *PLoS Neglected Tropical Diseases* 2018,
1153 **12:**e0006933.
1154 102. Hu B, Huang H, Wei Q, Ren M, Mburu DK, Tian X, Su J: **Transcription factors**
1155 **CncC/Maf and AhR/ARNT coordinately regulate the expression of multiple GSTs**
1156 **conferring resistance to chlorpyrifos and cypermethrin in *Spodoptera exigua*.** *Pest*
1157 *management science* 2019, **75:**2009-2019.
1158 103. Ingham VA, Pignatelli P, Moore JD, Wagstaff S, Ranson H: **The transcription factor**
1159 **Maf-S regulates metabolic resistance to insecticides in the malaria vector *Anopheles***
1160 ***gambiae*.** *BMC genomics* 2017, **18:**669.

104. Wilding CS: **Regulating resistance: CncC: Maf, antioxidant response elements and the overexpression of detoxification genes in insecticide resistance.** *Current opinion in insect science* 2018, **27:**89-96.

105. Kalsi M, Palli SR: **Transcription factors, CncC and Maf, regulate expression of CYP6BQ genes responsible for deltamethrin resistance in *Tribolium castaneum*.** *Insect Biochemistry and Molecular Biology* 2015, **65:**47-56.

106. Wittkopp PJ, Kalay G: **Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence.** *Nature Reviews Genetics* 2012, **13:**59-69.

107. Verta J-P, Jones FC: **Predominance of cis-regulatory changes in parallel expression divergence of sticklebacks.** *Elife* 2019, **8:**e43785.

108. Ingham V, Wagstaff S, Ranson H: **Transcriptomic meta-signatures identified in *Anopheles gambiae* populations reveal previously undetected insecticide resistance mechanisms.** *Nature communications* 2018, **9:**5282.

109. Clements J, Schoville S, Peterson N, Huseth AS, Lan Q, Groves RL: **RNA interference of three up-regulated transcripts associated with insecticide resistance in an imidacloprid resistant population of *Leptinotarsa decemlineata*.** *Pesticide Biochemistry and Physiology* 2016.

110. Kalsi M, Palli SR: **Transcription factor cap n collar C regulates multiple cytochrome P450 genes conferring adaptation to potato plant allelochemicals and resistance to imidacloprid in Leptinotarsa decemlineata (Say).** *Insect biochemistry and molecular biology* 2017, **83:**1-12.

111. Wilding CS: **Regulating resistance: CncC: Maf, antioxidant response elements and the overexpression of detoxification genes in insecticide resistance.** *Current Opinion in Insect Science* 2018.

112. Wray GA: **The evolutionary significance of cis-regulatory mutations.** *Nature Reviews Genetics* 2007, **8:**206-216.

113. Signor SA, Nuzhdin SV: **The evolution of gene expression in cis and trans.** *Trends in Genetics* 2018, **34:**532-544.

114. Roush RT: **Designing resistance management programs: how can you choose?** *Pesticide Science* 1989, **26:**423-441.

115. Denholm I, Rowland M: **Tactics for managing pesticide resistance in arthropods: theory and practice.** *Annual review of entomology* 1992, **37:**91-112.

116. Tabashnik BE: **Modeling and evaluation of resistance management tactics.** In *Pesticide Resistance in Arthropods.* Edited by Roush RT, Tabashnik BE. New York: Chapman & Hall; 1990: 153-182

117. Sudo M, Takahashi D, Andow DA, Suzuki Y, Yamanaka T: **Optimal management strategy of insecticide resistance under various insect life histories: Heterogeneous timing of selection and interpatch dispersal.** *Evolutionary applications* 2018, **11:**271-283.

118. Haridas C, Tenhumberg B: **Modeling effects of ecological factors on evolution of polygenic pesticide resistance.** *Journal of Theoretical Biology* 2018, **In Press**.

119. Alyokhin A, Mota-Sanchez D, Baker M, Snyder WE, Menasha S, Whalon M, Dively G, Moarsi WF: **The Red Queen in a potato field: integrated pest management versus chemical dependency in Colorado potato beetle control.** *Pest management science* 2015, **71:**343-356.

120. Margus A, Piiroinen S, Lehmann P, Tikka S, Karvanen J, Lindström L: **Sublethal pyrethroid insecticide exposure carries positive fitness effects over generations in a pest insect.** *Scientific reports* 2019, **9:**1-10.

121. Brevik K, Bueno EM, McKay S, Schoville SD, Chen YH: **Insecticide exposure affects intergenerational patterns of DNA methylation in the Colorado potato beetle, *Leptinotarsa decemlineata*.** *Evolutionary Applications* 2020.

122. Vontas J, Katsavou E, Mavridis K: **Cytochrome P450-based metabolic insecticide resistance in *Anopheles* and *Aedes* mosquito vectors: Muddying the waters.** *Pesticide biochemistry and physiology* 2020**:**104666.

123. Scott MJ, Gould F, Lorenzen M, Grubbs N, Edwards O, O'Brochta D: **Agricultural production: assessment of the potential use of Cas9-mediated gene drive systems for agricultural pest control.** *Journal of Responsible Innovation* 2018, **5:**S98-S120.

124. Gui S, Taning CNT, Wei D, Smagghe G: **First report on CRISPR/Cas9-targeted mutagenesis in the Colorado potato beetle, *Leptinotarsa decemlineata*.** *Journal of insect physiology* 2020, **121:**104013.

125. Palli SR: **RNA interference in Colorado potato beetle: steps toward development of dsRNA as a commercial insecticide.** *Current opinion in insect science* 2014, **6:**1-8.

126. Zhu KY, Palli SR: **Mechanisms, applications, and challenges of insect RNA interference.** *Annual review of entomology* 2019, **65:**2020.

127. He W, Xu W, Fu K, Guo W, Zhang J: **Low mismatch rate between double-stranded RNA and target mRNA does not affect RNA interference efficiency in Colorado potato beetle.** *Insects* 2020, **11:**449.

128. Yoon J-S, Mogilicherla K, Gurusamy D, Chen X, Chereddy SC, Palli SR: **Double-stranded RNA binding protein, Staufen, is required for the initiation of RNAi in coleopteran insects.** *Proceedings of the National Academy of Sciences* 2018, **115:**8334-8339.

129. Cappelle K, de Oliveira CFR, Van Eynde B, Christiaens O, Smagghe G: **The involvement of clathrin-mediated endocytosis and two Sid-1-like transmembrane proteins in double-stranded RNA uptake in the Colorado potato beetle midgut.** *Insect molecular biology* 2016, **25:**315-323.

130. Mehlhorn SG, Geibel S, Bucher G, Nauen R: **Profiling of RNAi sensitivity after foliar dsRNA exposure in different European populations of Colorado potato beetle reveals a robust response with minor variability.** *Pesticide Biochemistry and Physiology* 2020**:**104569.

131. Davis-Vogel C, Ortiz A, Procyk L, Robeson J, Kassa A, Wang Y, Huang E, Walker C, Sethi A, Nelson ME, Sashital DG: **Knockdown of RNA interference pathway genes impacts the fitness of western corn rootworm.** *Scientific Reports* 2018, **8:**7858.

132. Spit J, Philips A, Wynant N, Santos D, Plaetinck G, Broeck JV: **Knockdown of nuclease activity in the gut enhances RNAi efficiency in the Colorado potato beetle, *Leptinotarsa decemlineata*, but not in the desert locust, *Schistocerca gregaria*.** *Insect biochemistry and molecular biology* 2017, **81:**103-116.

133. Guan R-B, Li H-C, Fan Y-J, Hu S-R, Christiaens O, Smagghe G, Miao X-X: **A nuclease specific to lepidopteran insects suppresses RNAi.** *Journal of Biological Chemistry* 2018, **293:**6011-6021.

134. Luck RF, van den Bosch R, Garcia R: **Chemical insect control—a troubled pest management strategy.** *BioScience* 1977, **27:**606-611.

135. Chen YH, Schoville SD: **Editorial overview: Ecology: Ecological adaptation in agroecosystems: novel opportunities to integrate evolutionary biology and agricultural entomology.** 2018.

136. Nyerges Á, Csörgő B, Draskovits G, Kintses B, Szili P, Ferenc G, Révész T, Ari E, Nagy I, Bálint B: **Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance.** *Proceedings of the National Academy of Sciences* 2018, **115:**E5726-E5735.

137. Nevado B, Ramos-Onsins S, Perez-Enciso M: **Resequencing studies of nonmodel organisms using closely related reference genomes: optimal experimental designs and bioinformatics approaches for population genomics.** *Molecular ecology* 2014, **23:**1764-1779.

138. Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K, Wang J: **SNP detection for massively parallel whole-genome resequencing.** *Genome Research* 2009, **19:**1124-1132.

139. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2009, **25:**1754-1760.

140. R Core Team: **R: A language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing; 2019.

141. Wickham H: *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag; 2016.

142. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al: **The variant call format and VCFtools.** *Bioinformatics* 2011, **27:**2156–2158.

143. Lee TH, Guo H, Wang X, Kim C, Paterson AH: **SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data.** *BMC Genomics* 2014, **15:**162.

144. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Research* 2004, **32:**1792-1797.

145. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6.** University of Washington, Seattle: Distributed by the author. Department of Genome Sciences; 2005.

146. Luu K, Bazin E, Blum MG: **pcadapt: an R package to perform genome scans for selection based on principal component analysis.** *Molecular ecology resources* 2017, **17:**67-77.

147. Duforet-Frebourg N, Luu K, Laval G, Bazin E, Blum MG: **Detecting genomic signatures of natural selection with principal component analysis: Application to the 1000 Genomes Data.** *Molecular Biology and Evolution* 2016, **33:**1082-1093.

148. Cattell RB: **The scree test for the number of factors.** *Multivariate behavioral research* 1966, **1:**245-276.

149. Frichot E, François O: **LEA: an R package for landscape and ecological association studies.** *Methods in Ecology and Evolution* 2015, **6:**925-929.

150. Malinsky M: **Dsuite - fast D-statistics and related admixture evidence from VCF files.** *bioRxiv* 2019**:**634477.

151. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai WW, Fritz MHY, et al: **A Draft Sequence of the Neandertal Genome.** *Science* 2010, **328:**710-722.

152. Liu X, Fu Y-X: **Exploring population size changes using SNP frequency spectra.** *Nat Genet* 2015, **47:**555-559.

153. Terhorst J, Kamm JA, Song YS: **Robust and scalable inference of population history from hundreds of unphased whole genomes.** *Nature Genetics* 2017, **49:**303-312.

154. Allio R, Donega S, Galtier N, Nabholz B: **Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker.** *Molecular Biology and Evolution* 2017, **34:**2762–2772.

155. Keightley PD, Pinharanda A, Ness RW, Simpson F, Dasmahapatra KK, Mallet J, Davey JW, Jiggins CD: **Estimation of the spontaneous mutation rate in *Heliconius melpomene*.** *Molecular Biology and Evolution* 2014, **32:**239-243.

156. Oppold A-M, Pfenninger M: **Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in *Chironomus riparius*.** *Evolution Letters* 2017, **1-2:**86–92.

157. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD: **Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data.** *PLOS Genetics* 2009, **5:**e1000695.

158. Nadachowska-Brzyska K, Burri R, Smeds L, Ellegren H: **PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white Ficedula flycatchers.** *Molecular ecology* 2016, **25:**1058-1072.

159. Frichot E, Schoville S, Bouchard G, François O: **Testing for associations between loci and environmental gradients using latent factor mixed models.** *Molecular Biology and Evolution* 2013, **30:**1687-1699.

160. Crossley MS, Burke KD, Schoville SD, Radeloff VC: **Recent collapse of crop belts and declining diversity of US agriculture since 1840.** *Global Change Biology* 2021, **27:**151-164.

161. Bivand R, Keitt T, Rowlingson B, Pebesma E: ***rgdal*: bindings for the geospatial data abstraction library.** *R package version 08-16* 2014.

162. Hijmans R, van Etten J, Cheng J, Mattiuzzi M, Sumner M, Jonathan A: ***raster*: geographic data analysis and modeling.** 2017.

163. Crossley MS, Rondon SI, Schoville SD: **A comparison of resistance to imidacloprid in Colorado potato beetle (*Leptinotarsa decemlineata* Say) populations collected in the Northwest and Midwest US.** *American Journal of Potato Research* 2018**:**1-9.

164. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21:**3674-3676.

165. Feyereisen R, Dermauw W, Van Leeuwen T: **Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods.** *Pesticide Biochemistry and Physiology* 2015, **121:**61-77.

166. François O, Martins H, Caye K, Schoville Sean D: **Controlling false discoveries in genome scans for selection.** *Molecular Ecology* 2016, **25:**454-469.

167. De Leeuw CA, Neale BM, Heskes T, Posthuma D: **The statistical properties of gene-set analysis.** *Nature Reviews Genetics* 2016, **17:**353.

168. Rivals I, Personnaz L, Taing L, Potier M-C: **Enrichment or depletion of a GO category within a class of genes: which test?** *Bioinformatics* 2007, **23:**401-407.

169. Supek F, Bošnjak M, Škunca N, Šmuc T: **REVIGO summarizes and visualizes long lists of gene ontology terms.** *PloS one* 2011, **6:**e21800.

170. Kuznetsova I, Lugmayr A, Siira SJ, Rackham O, Filipovska A: **CirGO: an alternative circular way of visualising gene ontology terms.** *BMC bioinformatics* 2019, **20:**1-7.

171. Everaert C, Luypaert M, Maag JL, Cheng QX, Dinger ME, Hellemans J, Mestdagh P: **Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data.** *Scientific reports* 2017, **7:**1559.

172. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak MW, Gaffney DJ, Elo LL, Zhang X: **A survey of best practices for RNA-seq data analysis.** *Genome biology* 2016, **17:**13.

173. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements.** *Nature methods* 2015, **12:**357.

174. Liao Y, Smyth GK, Shi W: **The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads.** *Nucleic acids research* 2019, **47:**e47-e47.

175. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome biology* 2014, **15:**550.

176. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *Journal of the royal statistical society Series B (Methodological)* 1995, **57:**289-300.

177. Kolde R, Kolde MR: **Package 'pheatmap'.** *R Package* 2015, **1:**790.

1360

## Figure Legends

**Figure 1. A)** Unrooted phylogenetic tree of *Leptinotarsa* species obtained with SNPhylo, based on 35,838 SNPs (after LD-based reduction). Node labels represent bootstrap values. The blue arrow highlights a monophyletic clade comprising CPB samples collected in the U.S. and Europe. (S): imidacloprid susceptible population; (R): imidacloprid resistant population. **B)** Geographical sampling of *Leptinotarsa decemlineata* and estimated admixture coefficients. Admixture proportions were estimated with *SNMF* on the intergenic "CPB" dataset for k=6 clusters, and are shown as both pie charts and an individual bar-plot. Each pie chart represents a sampled location (small charts for single samples; large ones for populations of five individuals), referenced as a number. Colored boxes around large pie charts differentiate susceptible (green) vs. resistant samples (red). **C)** Population demographic histories (median Ne only) estimated from SMC++. Colors correspond to geographical regions.

**Figure 2**. Heatmap of *D*-statistics, showing the introgression patterns among CPB populations. The color of the heatmap cell indicates the most significant $D_{min}$ found with every population pairs: red colors indicate higher *D*-statistics, and generally more saturated colors indicate higher *P*-values. The complete biallelic dataset was analyzed.

**Figure 3**. Population tree showing the distribution of 24 resistance-associated selection events identified with hapFLK in the first 95 genomic scaffolds. Colors refer to geographical location. Internal branches show few selection events (one or two events in four branches, no selection event in five branches).

**Figure 4**. Gene expression heatmap among four populations of CPB larvae, showing divergent constitutive expression of 84 differentially expressed candidate insecticide resistance genes. Colors of expression levels correspond to log-fold change. See **Table S15** for the functional annotation of these genes.

1389 **Table 1**. Predictions from alternative mechanisms of rapid evolution to insecticide resistance in

1390 Colorado potato beetle.

| Evolutionary mechanism | Geographical pattern in resistant populations | Genome scans of selection | Haplotype-based selection scan | Differential gene expression |
|---|---|---|---|---|
| *De novo* mutation | Independent hard selective sweeps | A few statistically significant candidate genes | Long haplotype blocks around selected loci | Differential gene expression limited to key pathways |
| Rapid regulatory evolution | Selection in key regulatory genes leading to repeated upregulation of resistance pathways | A few statistically significant regulatory genes | Long haplotype blocks around regulatory genes encoding transcription factors | Differential expression of key transcription factors and constitutive expression differences in insecticide resistance pathways |
| Standing genetic diversity | Repeated selection on candidate insecticide resistance genes | Numerous statistically significant candidate genes | Short haplotype blocks around selected loci | Multiple differentially expressed genes in the same molecular pathways and constitutive expression differences in insecticide resistance pathways |

1391

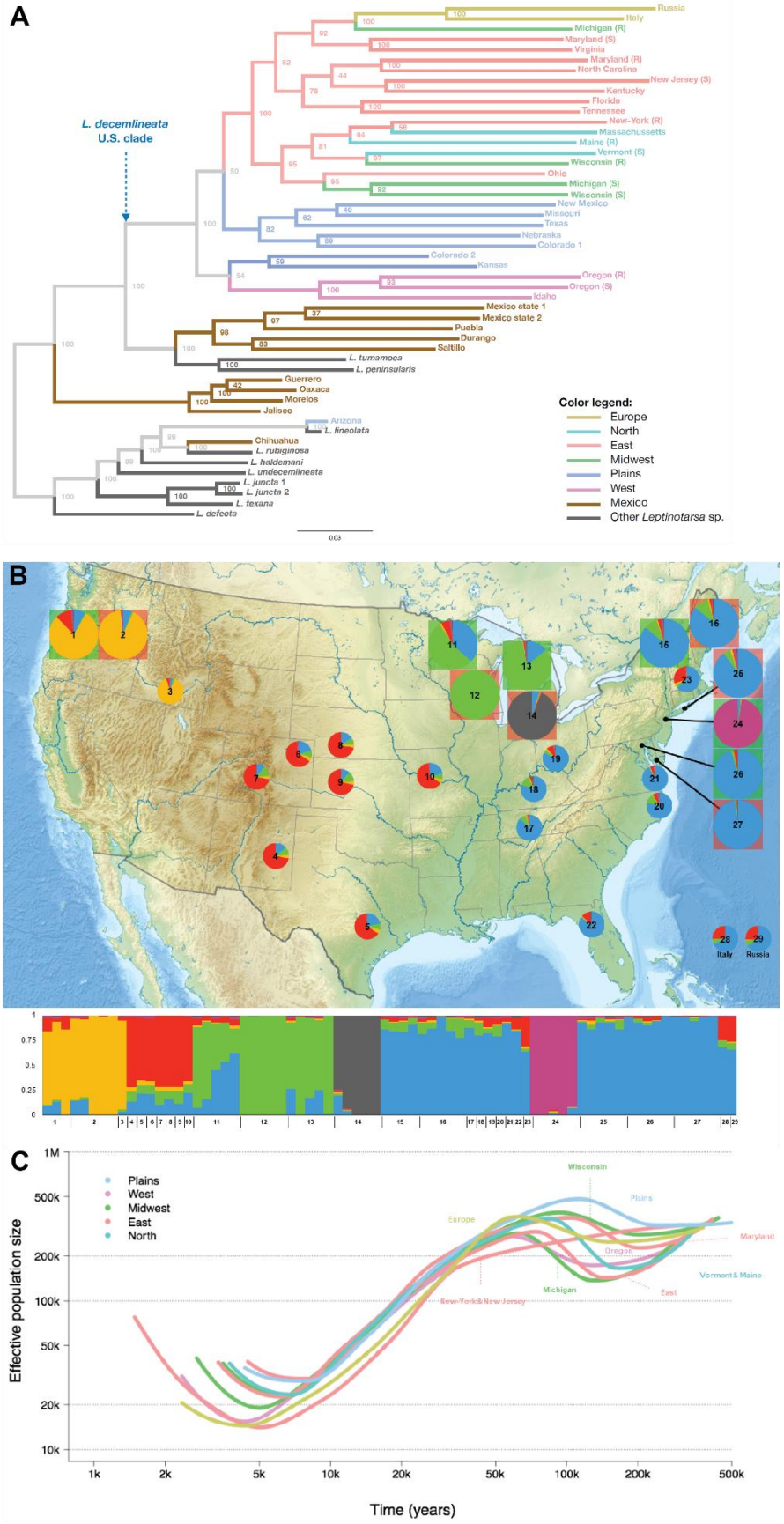
1392 **Table 2**. Candidate resistance genes identified in both PCAdapt and LFMM. The loading of each

1393 gene on a principal component is indicated (see **Additional File 1: Fig. S9**).

| Mechanisms | Categories | Gene ID* | Annotated Gene Name | Principal Component |
|---|---|---|---|---|
| Metabolic detoxification | ABC transporters | LDEC015007 | atp-binding cassette sub-family a member 5-like isoform x2 | 1 |
| | | LDEC002775 | atp-binding cassette sub-family g member 1-like | 4 |
| | | LDEC005530 | atp-binding cassette sub-family g member 4-like | 1 |
| | | LDEC019090 | Multidrug resistance-associated 1 | 1 |
| | | LDEC020530 | multidrug resistance-associated 4-like | 1 |
| | | LDEC022533 | multidrug resistance-associated protein 4 | 3 |
| | | LDEC005086 | multidrug resistance-associated protein 4 | 2 |
| | | LDEC002518 | probable multidrug resistance-associated protein lethal 03659 | 1 |
| | | LDEC012031 | probable multidrug resistance-associated protein lethal 03659 | 5 |
| | CYPs | LDEC018533 | cytochrome P450 | 3 |
| | | LDEC019188 | cytochrome P450 | 1 |
| | | LDEC019765 | cytochrome P450 | 1 |
| | | LDEC019766 | cytochrome P450 | 2 |
| | | LDEC015048 | cytochrome p450 | 4 |

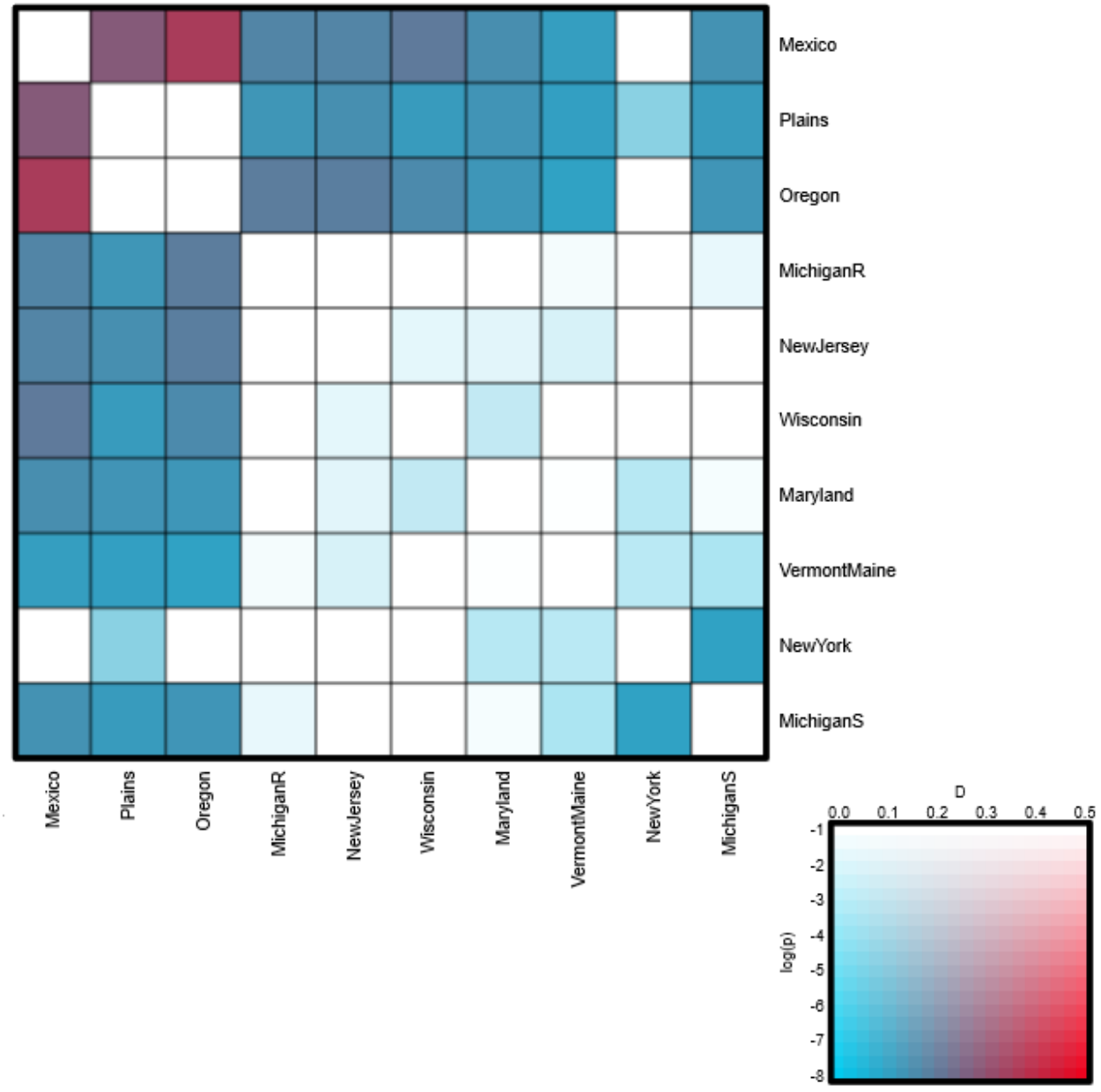
| | | | | |
|---|---|---|---|---|
| | | LDEC018119 | cytochrome p450 | 5 |
| | | LDEC005460 | cytochrome p450 6bq10 | 1 |
| | **Esterases** | LDEC017038 | esterase | 3 |
| | | LDEC018118 | esterase | 2 |
| | **GSTs** | LDEC012947 | glutathione s-transferase theta-1 | 1 |
| | **MFS** | LDEC009079 | major facilitator superfamily domain-containing protein 8 | 1 |
| **Target-sites** | **Voltage-dependent channels** | LDEC009862 | glutamate receptor 2-like isoform x4 | 1 |
| | | LDEC000112 | voltage-dependent calcium channel subunit alpha-2 delta-3 | 1 |
| | | LDEC021584 | Voltage-dependent calcium channel type D subunit alpha-1 | 1 |
| | | LDEC015955 | voltage-dependent calcium channel type d subunit alpha-1-like protein | 1 |
| | **Known insecticide resistance genes** | LDEC016101 | nicotinic acetylcholine receptor a9 subunit | 2 |
| | | LDEC007707 | nicotinic acetylcholine receptor subunit alpha4 | 1 |
| **Growth Factors** | **Cuticular proteins** | LDEC003392 | cuticle protein 19 | 2 |
| | | LDEC014693 | cuticular protein analogous to peritrophins 1-j precursor | 1 |
| | | LDEC010803 | larval cuticle protein 8-like | 3 |

*Among these genes, three (the CYP gene LDEC015048, the cuticular protein LDEC010803, and the voltage-dependent calcium channel gene LDEC000112) were found as candidate genes among field populations within Wisconsin [52]. The nicotinic acetylcholine receptor subunit alpha4 (LDEC007707) was also found as a candidate gene in a comparative genomic analysis of *Leptinotarsa* by Cohen *et al*. [67].




1397

**Figure 1. A)** Unrooted phylogenetic tree of *Leptinotarsa* species obtained with SNPhylo, based on 35,838 SNPs (after LD-based reduction). Node labels represent bootstrap values. The blue arrow highlights a monophyletic clade comprising CPB samples collected in the U.S. and Europe. (S): imidacloprid susceptible population; (R): imidacloprid resistant population. **B)** Geographical sampling of *Leptinotarsa decemlineata* and estimated admixture coefficients. Admixture proportions were estimated with *SNMF* on the intergenic "CPB" dataset for k=6 clusters, and are shown as both pie charts and an individual bar-plot. Each pie chart represents a sampled location (small charts for single samples; large ones for populations of five individuals), referenced as a number. Colored boxes around large pie charts differentiate susceptible (green) vs. resistant samples (red). **C)** Population demographic histories (median Ne only) estimated from SMC++. Colors correspond to geographical regions.




1409

1410 **Figure 2**. Heatmap of *D*-statistics, showing the introgression patterns among CPB populations.

1411 The color of the heatmap cell indicates the most significant $D_{min}$ found with every population

1412 pairs: red colors indicate higher *D*-statistics, and generally more saturated colors indicate higher

1413 *P*-values. The complete biallelic dataset was analyzed.

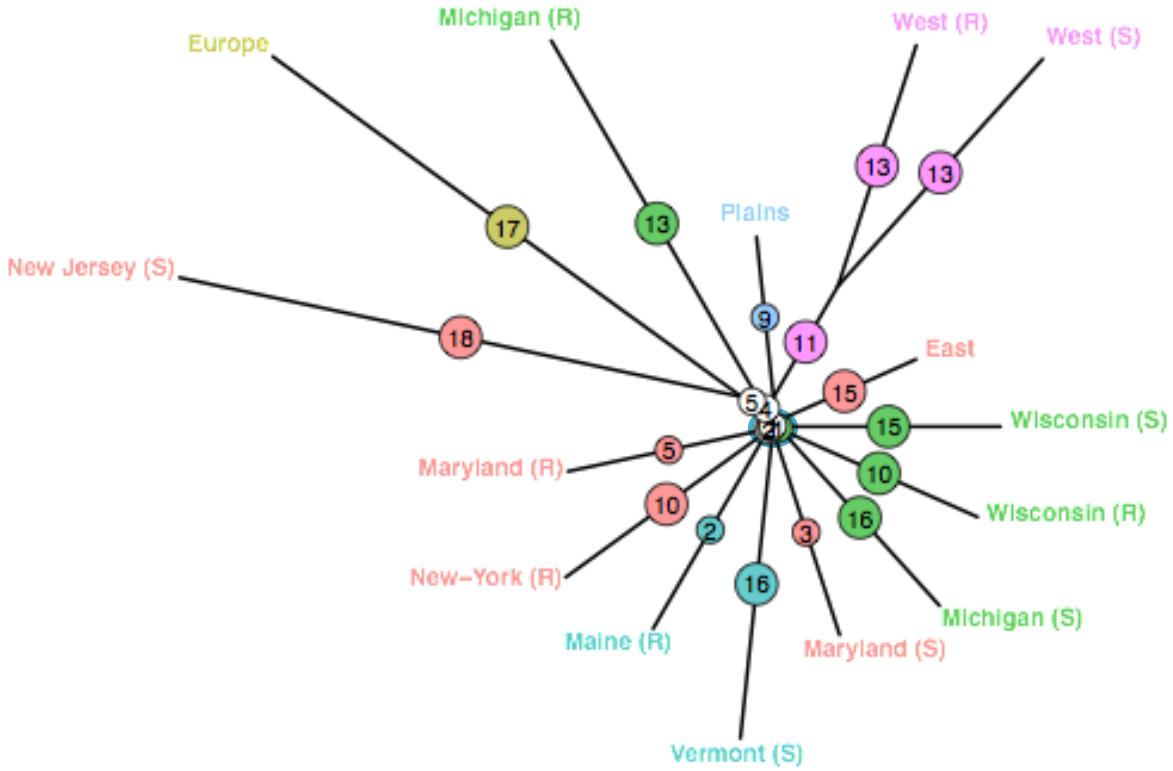



**Figure 3**. Population tree showing the distribution of 24 resistance-associated selection events identified with hapFLK in the first 95 genomic scaffolds. Most resistance-associated selection events occur on nine to eleven branches, and only two are singular to one branch. Colors refer to geographical location. Internal branches show few selection events (0.67 selection events on average).

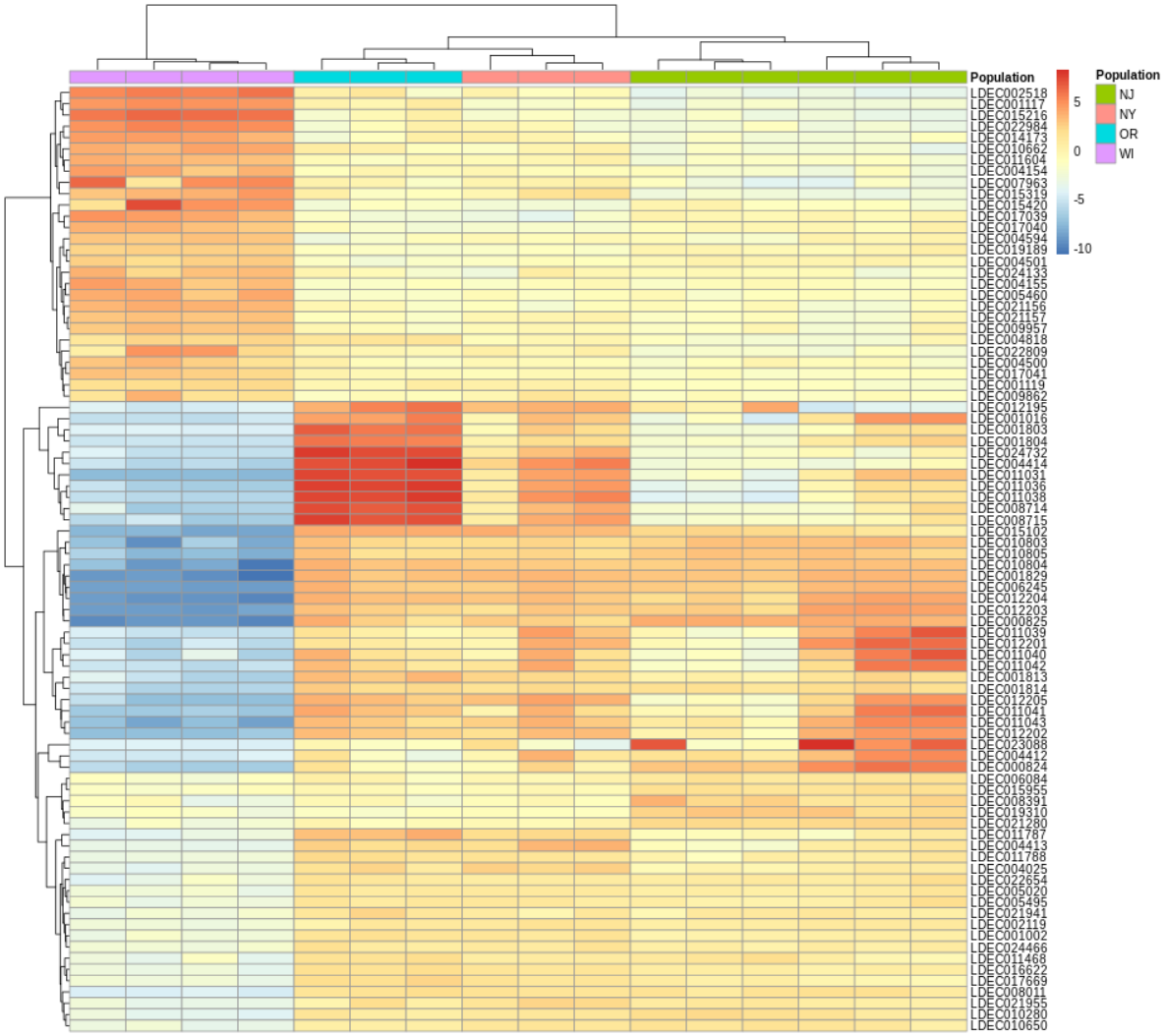



**Figure 4**. Gene expression heatmap among four populations of CPB larvae, showing divergent constitutive expression of 84 differentially expressed candidate insecticide resistance genes. Colors of expression levels correspond to log-fold change. See **Additional File 1: Table S12** for the functional annotation of these genes.