

## RESEARCH ARTICLE

## Genome mapping coupled with CRISPR gene editing reveals a P450 gene confers avermectin resistance in the beet armyworm

Yayun Zuo<sup>1,2</sup>, Yu Shi<sup>1</sup>, Feng Zhang<sup>1</sup>, Fang Guan<sup>1</sup>, Jianpeng Zhang<sup>1</sup>, René Feyereisen<sup>3</sup>, Jeffrey A. Fabrick<sup>4</sup>, Yihua Yang<sup>1\*</sup>, Yidong Wu<sup>1\*</sup>

**1** The Key Laboratory of Plant Immunity and College of Plant Protection, Nanjing Agricultural University, Nanjing, China, **2** Institute of Pesticide Science, College of Plant Protection, Northwest A&F University, Yangling, Shaanxi, China, **3** Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark, **4** USDA ARS, U.S. Arid Land Agricultural Research Center, Maricopa, Arizona, United States of America

\* [yhyang@njau.edu.cn](mailto:yhyang@njau.edu.cn) (YY); [wyd@njau.edu.cn](mailto:wyd@njau.edu.cn) (YW)

## OPEN ACCESS

**Citation:** Zuo Y, Shi Y, Zhang F, Guan F, Zhang J, Feyereisen R, et al. (2021) Genome mapping coupled with CRISPR gene editing reveals a P450 gene confers avermectin resistance in the beet armyworm. *PLoS Genet* 17(7): e1009680. <https://doi.org/10.1371/journal.pgen.1009680>

**Editor:** Subba Reddy Palli, University of Kentucky, UNITED STATES

**Received:** April 26, 2021

**Accepted:** June 23, 2021

**Published:** July 12, 2021

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** The project and raw sequencing data have been deposited at the NCBI under the accessions PRJNA588360 and SRR10441878-SRR10441889, and at the FigShare ([https://figshare.com/projects/Genome\\_assembly\\_and\\_annotations\\_of\\_the\\_Spodoptera\\_exigua\\_beet\\_armyworm\\_/100319](https://figshare.com/projects/Genome_assembly_and_annotations_of_the_Spodoptera_exigua_beet_armyworm_/100319)).

**Funding:** This work was funded by grants to YY from the National Natural Science Foundation of China (Grant No. 31572030) and National Science & Technology Fundamental Resources

## Abstract

The evolution of insecticide resistance represents a global constraint to agricultural production. Because of the extreme genetic diversity found in insects and the large numbers of genes involved in insecticide detoxification, better tools are needed to quickly identify and validate the involvement of putative resistance genes for improved monitoring, management, and countering of field-evolved insecticide resistance. The avermectins, emamectin benzoate (EB) and abamectin are relatively new pesticides with reduced environmental risk that target a wide number of insect pests, including the beet armyworm, *Spodoptera exigua*, an important global pest of many crops. Unfortunately, field resistance to avermectins recently evolved in the beet armyworm, threatening the sustainable use of this class of insecticides. Here, we report a high-quality chromosome-level assembly of the beet armyworm genome and use bulked segregant analysis (BSA) to identify the locus of avermectin resistance, which mapped on 15–16 Mbp of chromosome 17. Knockout of the *CYP9A186* gene that maps within this region by CRISPR/Cas9 gene editing fully restored EB susceptibility, implicating this gene in avermectin resistance. Heterologous expression and *in vitro* functional assays further confirm that a natural substitution (F116V) found in the substrate recognition site 1 (SRS1) of the CYP9A186 protein results in enhanced metabolism of EB and abamectin. Hence, the combined approach of coupling gene editing with BSA allows for the rapid identification of metabolic resistance genes responsible for insecticide resistance, which is critical for effective monitoring and adaptive management of insecticide resistance.

## Author summary

Insecticide resistance is a global constraint to agricultural production, and rapid identification of resistance genes is critical for better monitoring and management of resistant insect pests. Identification of metabolic resistance genes has always been a challenging

Investigation Program of China (Grant No. 2018FY101103). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

task due to the high diversity of insect detoxification enzyme genes. Here, we report a high-quality chromosome-level assembly of the beet armyworm genome and use bulked segregant analysis (BSA) to identify the locus of avermectin resistance, which mapped on 15–16 Mbp of chromosome 17. Knockout of the *CYP9A186* gene that maps within this region by CRISPR/Cas9 gene editing fully restored avermectin susceptibility. Heterologous expression and *in vitro* functional assays further confirm that a natural substitution (F116V) found in the substrate recognition site 1 (SRS1) of the CYP9A186 protein results in enhanced metabolism of avermectin. Hence, the combined approach of coupling gene editing with BSA allows for the rapid identification of metabolic resistance genes responsible for insecticide resistance, which is critical for field monitoring of such mutations, for making improved decisions on appropriate use of effective chemistries, as well for improvements in the design of future compounds that target *S. exigua*.

## Introduction

Avermectins are a group of 16-membered macrocyclic lactones possessing potent anthelmintic and insecticidal activities [1]. Emamectin benzoate (EB) [(4'R)-4"-deoxy-4"-(methylamino) avermectin B1 benzoate] is a macrocyclic lactone semi-synthetic derivative of the avermectin family and has much higher potency than abamectin against a number of lepidopterous pests [2]. Similar to abamectin and ivermectin, the mode of action of EB in invertebrates is through the overstimulation of glutamate-gated chloride channels (GluCl<sub>s</sub>) [1–3]. With extensive field application of avermectins, the evolution of resistance in mites, insects, and parasitic worms has become an increasingly serious problem in agriculture [4–7].

At least two major mechanisms are known to cause avermectin resistance in arthropods: target-site insensitivity and metabolic enzyme-based resistance [8]. Target site mutations in the glutamate-gated chloride channels are associated with resistance to avermectins in *Drosophila melanogaster* [9], *Plutella xylostella* [10,11] and *Tetranychus urticae* [5,12]. Enhanced oxidative metabolism of abamectin has also been reported as a resistance mechanism in *Leptinotarsa decemlineata* [13], *Bemisia tabaci* [14], *P. xylostella* [15] and *T. urticae* [16]. Enhanced avermectin metabolism due to increased expression of cytochrome P450 was shown in *T. urticae* [17] and in *Pediculus humanus humanus* [3]. Other detoxification mechanisms, including sequestration and/or the metabolic action of esterases, carboxylesterases, and/or glutathione S-transferases (GSTs) may also play a role in avermectin resistance in some arthropods [16,18]. However, the underlying molecular mechanisms of resistance to EB and abamectin remain largely unknown in lepidopteran pests.

The beet armyworm, *Spodoptera exigua*, is an important global pest of many crops [6,19] and because of repeated overuse of EB, it has evolved high levels of insecticide resistance in the field in some parts of Asia, including Pakistan and China [6,19]. Bulk segregant analysis (BSA) has recently been used to rapidly identify genomic regions associated with unique phenotypes in numerous different species [20]. To better understand the molecular genetic basis of EB resistance, we first produced a chromosome-level *de novo* genome assembly from a EB-susceptible reference beet armyworm laboratory strain from China. We then used BSA to identify loci containing major resistance genes and used CRISPR/Cas9 gene-editing to map resistance to the *CYP9A186* gene. Furthermore, the natural substitution (F116V) in the substrate binding site 1 (SRS1) region of CYP9A186 was functionally verified to confer resistance to both avermectins, EB and abamectin. Hence, the combined use gene editing with BSA provided a rapid and highly efficient means to identify metabolic resistance genes responsible for

insecticide resistance. Such results are critical for the monitoring of field populations of the beet armyworm for avermectin resistance and for the future use and/or design of new chemistries to manage resistant populations.

## Results

### Chromosome-level genome assembly

To facilitate genetic mapping of resistance, we first sequenced the beet armyworm genome using both Illumina and PacBio whole genome sequencing (WGS). We generated 3,919,515 PacBio subreads yielding 38.98 Gb (87X coverage) and 56.32 Gb of Hi-C sequencing data (126X coverage), respectively. Estimated genome size ranged from 408.58 Mb to 448.90 Mb with a low heterozygosity for the sequenced susceptible WH-S strain (S1 Table). The assembled genome consisted of 667 contigs spanning 446.80 Mb with a scaffold/contig N50 length of 14.36/3.47 Mb (Table 1); among them, 367 contigs were anchored into 32 pseudo-chromosomes (Fig 1), accounting for greater than 96% (429.74 Mb) of the genome. Consistent with lepidopteran females being heterogametic for sex chromosomes [21], 30 autosomes plus two sex chromosomes (Z and W) were identified from the sequenced female pupa. When compared with the insect\_odb9 reference dataset consisting of 1,658 functional genes, BUSCO completeness analyses showed our final genome assembly was 97.9% complete, with only 0.2% fragmented and 1.9% missing BUSCO genes (S2 Table). Reflecting the high degree of completeness and accuracy of the genome assembly, we mapped 97.62% Illumina short reads and 93.78% of the PacBio long reads to our final assembly with very low redundancy (2.1% duplicated BUSCOs) (Table 2). Among four Noctuidae pest species (Table 1), assemblies of *S. exigua* and *S. frugiperda* have the highest genome contiguity (contig N50 > 1 Mb, contig number < 1,000) and the lowest gap content (< 0.1%).

### Gene annotation

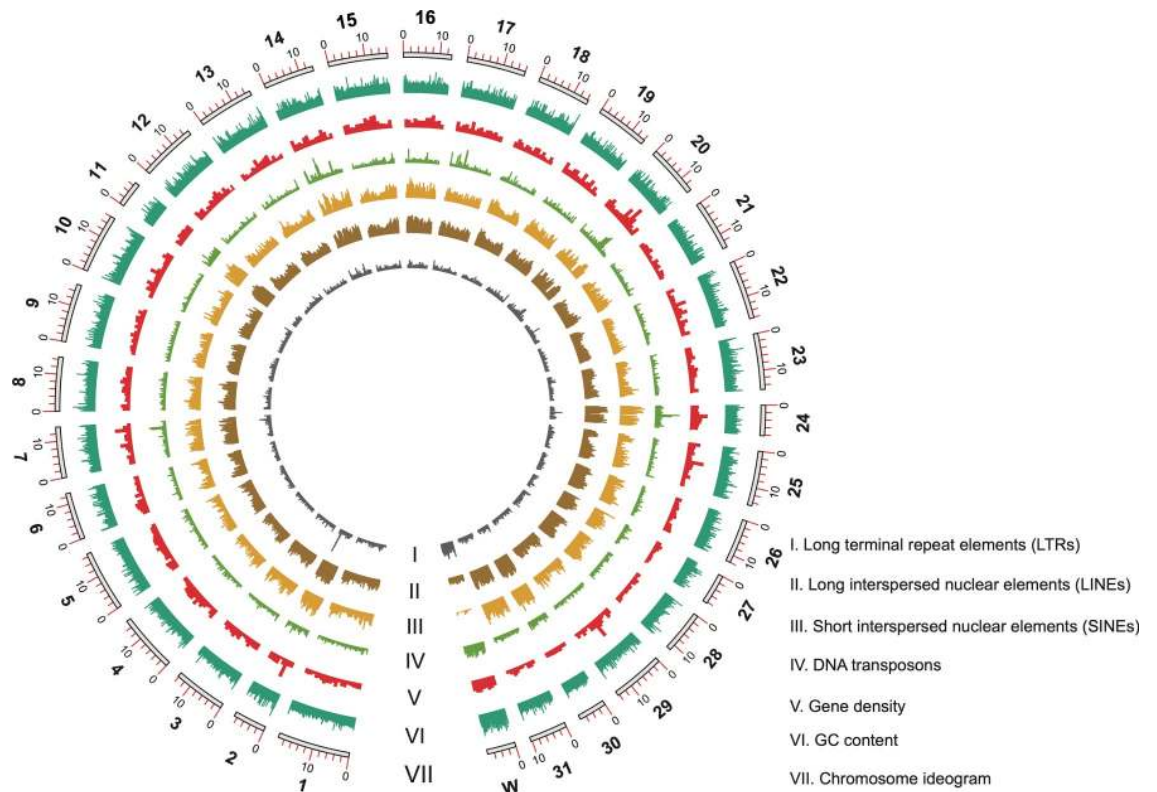
To comprehensively annotate genes in our assembled beet armyworm genome, we integrated *ab initio*, transcriptome- and protein homology-based strategies to predict 17,727 protein-coding genes (e.g., gene models) with an overall mean length of 9,605 bp (Table 2). BUSCO assessment of the gene models positively identified 1,601 of the 1,658 reference genes (96.6%). Of

**Table 1. Genome assembly metrics for four Noctuidae species.**

	<i>S. exigua</i>	<i>S. litura</i> <sup>a</sup>	<i>S. frugiperda</i> <sup>a</sup>	<i>H. armigera</i> <sup>a</sup>
Accession number	GCA_011316535.1	GCA_002706865.1	GCA_011064685.1	GCA_002156985.1
Assembly size (Mb)	446.80	438.96	486.29	337.09
Number of scaffolds	301	2,975	92	998
Longest scaffold (Mb)	19.74	18.94	22.37	6.15
Scaffold N50 length (Mb)	14.36	13.59	16.35	1.00
Number of contigs	667	13,637	617	24,552
Longest contig (Mb)	15.88	0.64	16.14	0.31
Contig N50 length (Mb)	3.47	0.068	1.13	0.023
Pseudo-chromosomes	32	31	32	31
GC (%)	36.67	36.56	36.43	37.50
Gaps (%)	0.075	2.49	0.011	11.01
BUSCO completeness (%)	97.9	97.1	93.4	96.8

<sup>a</sup> Reference sources: *Spodoptera litura* [22], *Spodoptera frugiperda* [87] and *Helicoverpa armigera* [88,89].

<https://doi.org/10.1371/journal.pgen.1009680.t001>



**Fig 1. Circos plot of beet armyworm genome.** Tracks from innermost to outermost are: I. Long terminal repeat elements (LTRs), II. Long interspersed nuclear elements (LINEs), III. Short interspersed nuclear elements (SINEs), IV. DNA transposons, V. Gene density, VI. GC content, and VII. Chromosome ideogram with numbers indicating the chromosome nomenclature order and size (in Mb). A 10 kb sliding window is shown for the GC content track and 100 kb for all other tracks.

<https://doi.org/10.1371/journal.pgen.1009680.g001>

these genes, only 45 (2.7%) were duplicated, 12 (0.7%) were fragmented, and 45 (2.7%) were missing, indicating the assemble genome is well-represented by its protein-coding genes. Comparison of the genome with RNA-seq data indicated a total mean exon and intron count of 6.4 and 5.2 per gene, with a mean length of 335 bp and 1,435 bp, respectively (Table 2). Among the predicted genes, 82% had transcriptome support and 96% had matches within the UniProt database. InterProScan identified protein domains for 13,186 (74%) genes, GO terms for 8,227 (47%) genes and Reactome pathways for 2,731 (15%) genes, respectively. EggNOG predicted GO terms for 6,706 genes, KEGG orthology (KO) matches for 7,643 genes, KEGG pathway matches for 4,493 genes, EC matches for 2,675 genes, and COG matches for 12,987 genes (Table 2).

Using *de novo* prediction and library alignments to detect repetitive DNA, we found that 33% (147.97 Mb) of the beet armyworm genome corresponds to putative transposable elements (TEs). This is consistent with the reported repeat content (32%) published for *S. litura* [22]. Several families of conserved TEs were found in the beet armyworm genome, including long interspersed nuclear elements (LINEs, 15%), rolling-circle transposition (RC, 4%), long terminal repeat elements (LTR, 3%), short interspersed nuclear elements (SINE, 3%), and DNA transposons (2%) (S3 Table). Simple repeats also occupied approximately 1% of the genome (S3 Table). Consistent with other lepidopterans [21], the beet armyworm chromosome W is enriched with DNA transposons and LTR retrotransposons when compared with the other 30 autosomes and the Z chromosome (Fig 1).

**Table 2. Genome annotation statistics of the beet armyworm.**

Elements	Number
Protein-coding genes	17,727
Mean protein (aa) and gene (bp) length	482.19/9,604.99
Exons/introns per gene	6.40/5.20
Exon	37.95 Mb (8.49%)
Intron	132.31 Mb (28.34%)
Mean exon/intron length	334.78/1,434.55
BUSCO completeness (%)	96.6
GO	10,298
Reactome pathway	2,731
KEGG orthology	7,643
KEGG pathway	4,493
EC	2,675
COG	12,987

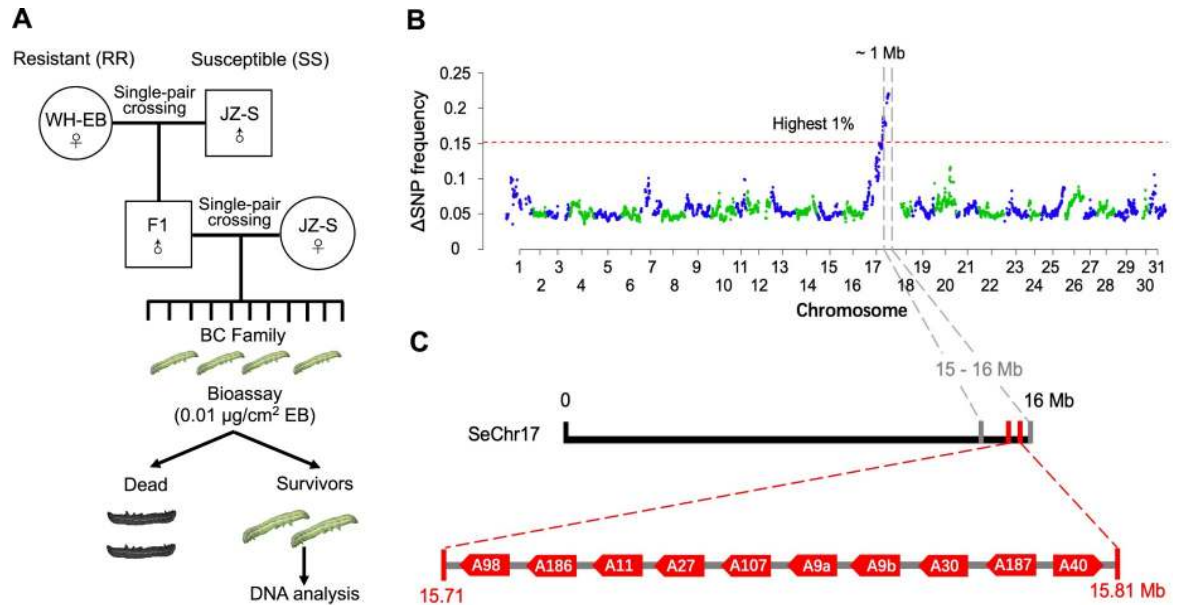
<https://doi.org/10.1371/journal.pgen.1009680.t002>

### Bulked segregant analysis

Tissue samples (120 moth legs) from EB-resistant individuals (survivors) derived from a backcrossed family generated by crossing a resistant WH-EB female with a susceptible JZ-S male (Fig 2A) were pooled and used for bulked sequence analysis (BSA). In doing so, we identified a 1 Mb genomic region in chromosome 17 (from 15–16 Mb) that was significantly biased with resistance (Fig 2B). A cluster (chromosomal locations: 15.71 Mb to 15.81Mb) of *CYP* genes belonging to the *CYP9A* family is located in this region of chromosome 17 (Fig 2C), whose functions are frequently associated with xenobiotic detoxification [22–25]. Detailed analysis of the *CYP9A* gene cluster from the genome of *S. exigua* identified ten *CYP9A* genes, which were named *CYP9A40*, *CYP9A187*, *CYP9A30*, *CYP9A9a* and *CYP9A9b* (duplication), *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98* according to the International P450 Nomenclature Committee (Fig 2C). We therefore hypothesize that the dominant EB resistance in WH-EB is associated with one or more P450 genes in the *CYP9A* gene cluster of WH-EB.

### Genetic mapping of resistance using CRISPR/Cas9

To map resistance in the WH-EB strain, we created four homozygous knockout strains (dA40-A98, dA40-A107, dA107-A98, and dA186) and observed their corresponding resistance phenotypes (Fig 3). The first knockout of the entire *CYP9A40* to *CYP9A98* gene cluster ( $\Delta$ *CYP9A98-CYP9A40*) resulted in a strain (e.g., dA40-A98) with complete restoration of susceptibility to EB (Table 3), indicating that at least one resistance gene is located in this region. Subsequent CRISPR/Cas9 knockout of the chromosomal fragment corresponding to *CYP9A40* to *CYP9A107* ( $\Delta$ *CYP9A107-CYP9A40*) resulted in a strain (dA40-A107) that remained fully resistant to both insecticides (Table 3). In contrast, susceptibility was fully restored in the dA107-A98 strain where the chromosome fragment from *CYP9A107* to *CYP9A98* ( $\Delta$ *CYP9A107-CYP9A98*) was deleted (Table 3), further narrowing the resistance gene to this region. Finally, we used CRISPR/Cas9 to generate a strain (dA186) homozygous for a 4-bp deletion in *CYP9A186* ( $\Delta$ *CYP9A186*) that results in the introduction of a premature stop codon and the loss of downstream active sites within the cytochrome P450 enzyme (S1 Fig). Bioassays show that susceptibility to abamectin and EB was fully restored in the dA186 strain (Table 3 and Fig 3), confirming that resistance maps to the *CYP9A186* gene.



**Fig 2. Experimental design for crosses, bioassays and bulked segregant analysis (BSA) to map emamectin benzoate resistance.** (A) Single-pair crosses between JZ-S and WH-EB virgin adults produced families of hybrid F<sub>1</sub> offspring. A male F<sub>1</sub> was backcrossed to a female JZ-S to produce the backcross (BC) family. Progeny from the BC family were exposed to a diagnostic concentration of EB of which 120 survivors and their F<sub>1</sub> parents were analyzed by next-generation sequencing (NGS). (B) BSA mapping revealed a 1 Mb genomic region in chromosome 17 highly correlated with EB resistance in WH-EB. The mean frequency deviation values obtained from a sliding window analysis (1 Mb window with 100 Kb step) were plotted across all 31 chromosomes of the genome. The red dashed line corresponds to the top 1% threshold of the SNP-index likely containing changes linked with EB resistance. (C) Expanded map of chromosome 17 (from 15 Mb to 16 Mb) that includes the *CYP9A* gene cluster (15.71 Mb to 15.81 Mb).

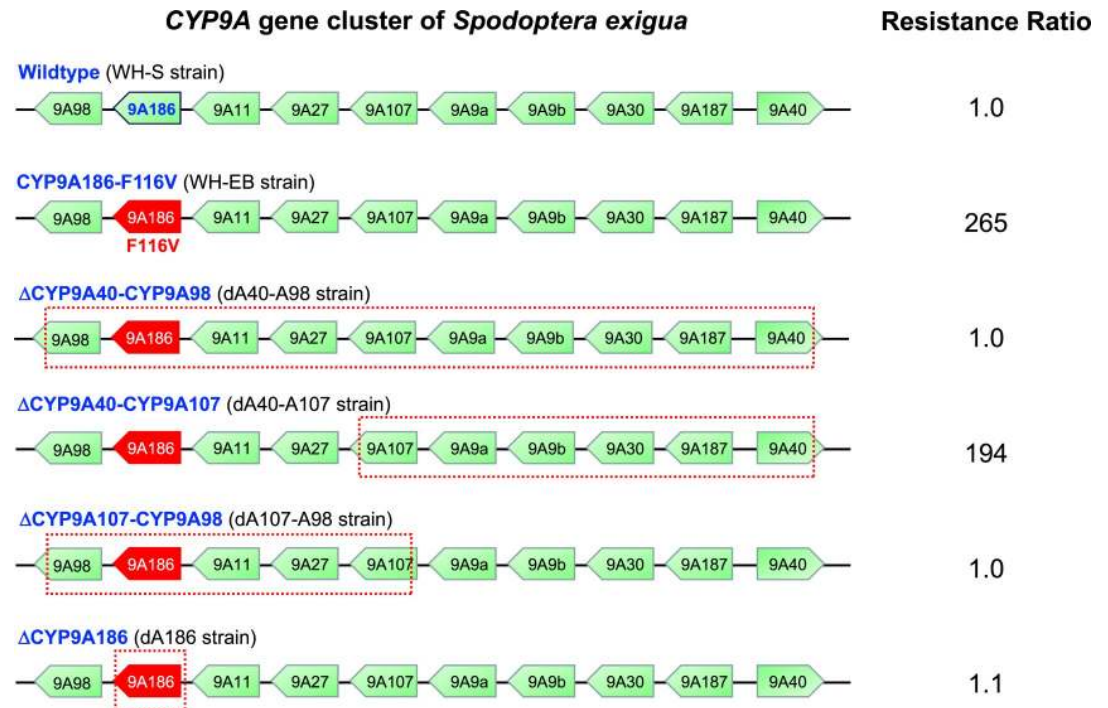
<https://doi.org/10.1371/journal.pgen.1009680.g002>

### Identification of a natural mutation in *CYP9A186* from WH-EB associated with avermectin resistance

To determine if one of the P450 genes in the defined region associated with avermectin resistance differed between susceptible and resistant strains, we PCR-amplified, cloned and sequenced the full-length cDNAs corresponding to the open reading frames (ORF) of *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98* from WH-S and WH-EB, respectively. We found a single T346G mutation resulting in the F116V substitution within the SRS1 region of *CYP9A186* (GenBank accession no. MN179472) in WH-EB and not WH-S (Fig 4A). DNA based genotyping of 20 larvae from each strain further revealed that WH-EB is homozygous for Val (GTT) at codon 116, whereas WH-S is homozygous for Phe (TTT) (Fig 4B). No other consistent mutations were found in the other four genes. Additionally, we also observed a 10-fold increase in *CYP9A186* transcript abundance (i.e., overexpression) in fat body tissue from WH-EB compared with WH-S (Fig 4C). None of the other four P450 candidate genes showed higher transcription in WH-EB compared to WH-S (Fig 4C). These data therefore show that a naturally-occurring point mutation and/or overexpression of *CYP9A186* is associated with avermectin resistance in WH-EB strain and provides additional supports for its involvement in metabolic insecticide resistance in the beet armyworm.

### Heterologous expression of P450s in cultured insect cells

Both wild-type (*CYP9A186*<sup>wt</sup>) and mutant (*CYP9A186*-F116V) *CYP9A186* enzymes were expressed in *Trichoplusia ni* High Five (Hi5) cells using a baculovirus system. Reduced CO-diffusion spectra of *CYP9A186*<sup>wt</sup> and *CYP9A186*-F116V in purified recombinant microsomes



**Fig 3. Genome mapping of the CYP9A gene cluster by CRISPR/Cas9 gene editing.** Deletions in the CYP9A gene cluster are indicated with dotted boxes and include  $\Delta$ CYP9A40-CYP9A98,  $\Delta$ CYP9A40-CYP9A107,  $\Delta$ CYP9A107-CYP9A98, and  $\Delta$ CYP9A186. Resistance ratios to EB for each of the homozygous strains harboring both the natural CYP9A186-F116V mutation in resistant WH-EB strain and the CRISPR/Cas9-edited strains are shown.

<https://doi.org/10.1371/journal.pgen.1009680.g003>

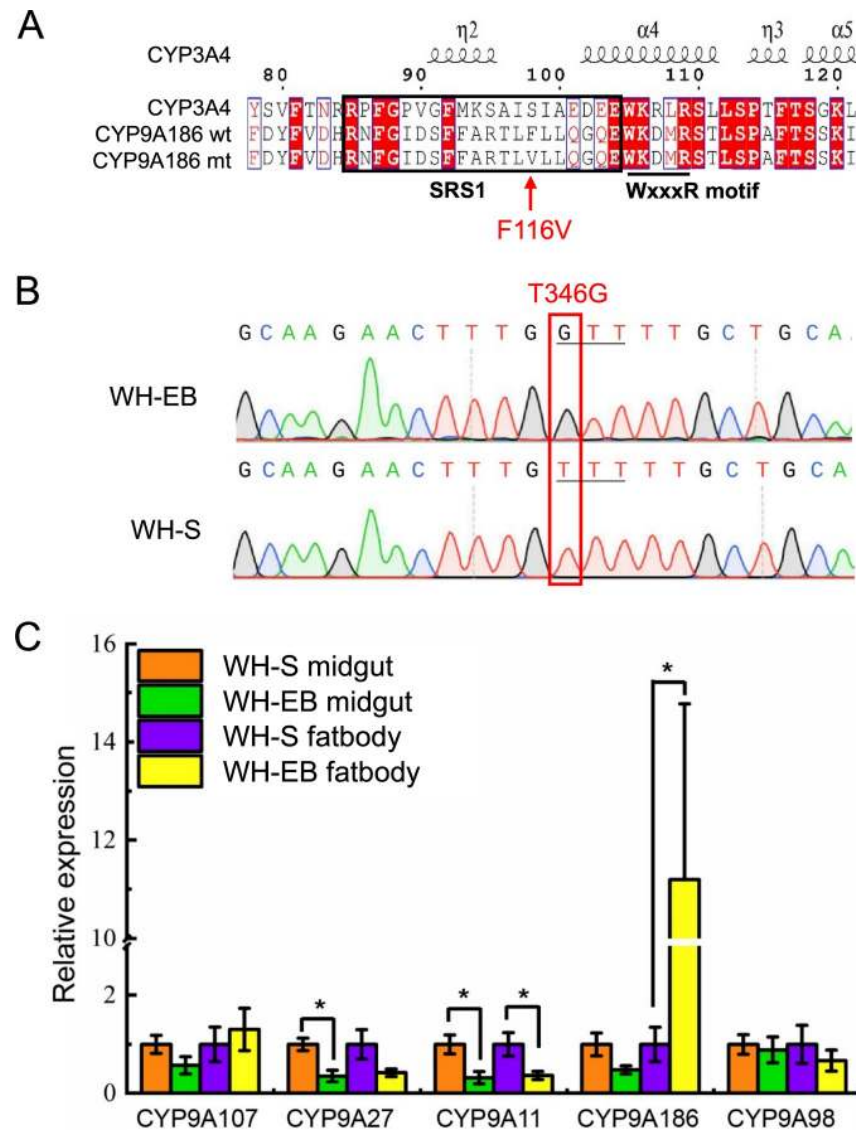
showed a maximum peak near 450 nm (S2 Fig), indicating successful expression of functional enzymes within the Hi5 cells. In addition, both recombinant P450s showed O-debenzylation activities to 7-benzyloxy-4-trifluoromethyl coumarin (BFCOD), again indicating that the recombinant enzymes had correct folding and were active. Microsomes with CYP9A186wt and CYP9A186-F116V had BFCOD specific activities of  $0.083 \pm 0.001$  and  $0.057 \pm 0.003$

**Table 3. Response to abamectin and emamectin benzoate of beet armyworm larvae from a susceptible strain (WH-S), a resistant strain (WH-EB), and four CRISPR/Cas9 gene-edited CYP9A knockout strains (dA40-A98, dA40-A107, dA107-A98, and dA186).**

Insecticide	Strain	Slope	LC <sub>50</sub> (μg/cm <sup>2</sup> )	95% Fiducial limits	RR <sup>a</sup>
Abamectin	WH-S	3.06±0.44	0.31	0.25–0.44	-
	WH-EB	2.60±0.36	13.89	10.86–19.10	45
	dA40-A98	2.53±0.34	0.26	0.20–0.35	0.8
	dA40-A107	2.26±0.33	15.72	11.93–22.93	51
	dA107-A98	4.23±0.66	0.45	0.37–0.56	1.4
	dA186	3.92±0.59	0.52	0.43–0.68	1.7
Emamectin benzoate	WH-S	2.57±0.35	0.0017	0.0014–0.0022	-
	WH-EB	3.39±0.48	0.45	0.37–0.58	265
	dA40-A98	3.33±0.51	0.0018	0.0014–0.0023	1.0
	dA40-A107	2.60±0.33	0.33	0.26–0.43	194
	dA107-A98	3.37±0.56	0.0017	0.0014–0.0021	1.0
	dA186	2.46±0.33	0.0019	0.0015–0.0024	1.1

<sup>a</sup> Resistance ratio = LC<sub>50</sub> (WH-EB or knockouts) / LC<sub>50</sub> (WH-S).

<https://doi.org/10.1371/journal.pgen.1009680.t003>



**Fig 4. Identification of the *CYP9A186* T346G mutation resulting in the F116V substitution and quantitative real-time PCR transcriptional analysis of *CYP9A186*.** (A) Alignment of the deduced *CYP9A186* amino acid sequences from WH-EB and WH-S with a translated human cytochrome P450 (e.g., CYP3A4). The F116V substitution is shown with a red arrow. The predicted locations of alpha helical secondary structures are marked based on the template of the CYP3A4 protein structure (1TQN) using ESPrpt 3.0. (B) Representative chromatograms from direct sequencing of *CYP9A186* PCR products from WH-EB and WH-S showing the T346G mutation. (C) qRT-PCR analysis of *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98* in midgut and fat body from WH-S and WH-EB. Data shown represent means  $\pm$  SE derived from four biological replicates. Asterisks indicate significant differences between the strains (Student's *t*-test,  $p < 0.05$ ).

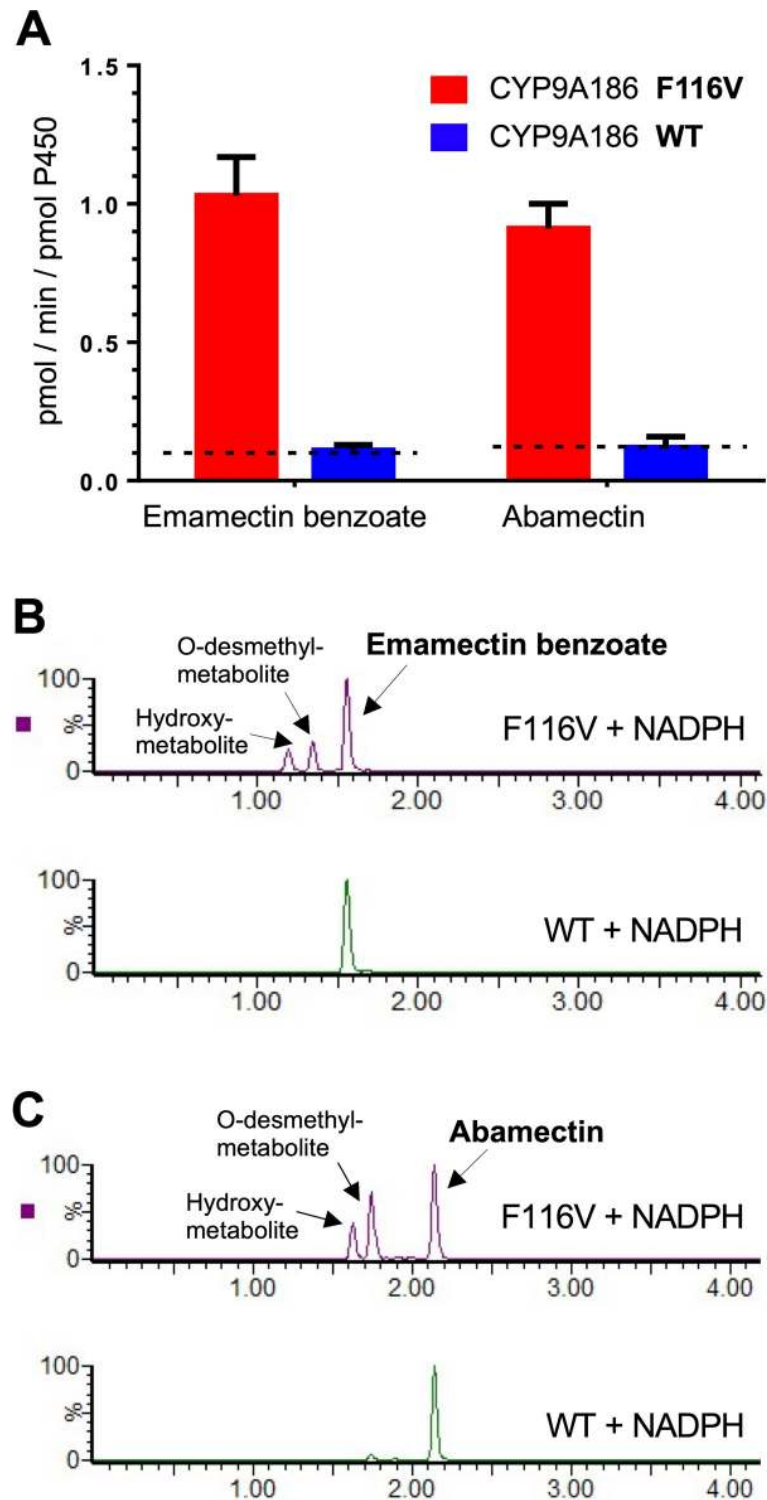
<https://doi.org/10.1371/journal.pgen.1009680.g004>

pmol/min/pmol P450, respectively (mean values  $\pm$  SEM,  $n = 4$ ). Non-insertion control Hi5 microsomes showed no BFCOD activity.

### EB and abamectin metabolism and identification of their metabolites

EB, abamectin and their multiple metabolites showed excellent separation and linearity under our experimental conditions (Figs 5 and S3). Recovery rates for both EB and abamectin were nearly 100% for the non-insertion control samples. Our *in vitro* assays showed that EB and





**Fig 5. Metabolism of emamectin benzoate and abamectin by microsomes containing recombinant CYP9A186 and CYP9A186-F116V.** (A) Specific activity (pmol/min/pmol P450) of EB and abamectin hydrolysis. Data are mean values  $\pm$  SEM ( $n = 6$ ). Dotted lines show mean limits of detection (LOD). There is no significant difference between WT and LOD (One sample  $t$ -test,  $p = 0.69$  and  $0.83$  for emamectin benzoate and abamectin, respectively). (B) MRM signals of emamectin benzoate and its metabolites. (C) MRM signals of abamectin and its metabolites.

<https://doi.org/10.1371/journal.pgen.1009680.g005>

abamectin were both metabolized by CYP9A186-F116V microsomal fractions at higher rates than by CYP9A186wt microsomes (Fig 5A). Whereas CYP9A186-F116V microsomes metabolized EB and abamectin at  $1.03 \pm 0.14$  pmol/min/pmol P450 and  $0.91 \pm 0.09$  pmol/min/pmol P450, respectively, the metabolic activity of both insecticides by CYP9A186wt microsomal fractions was negligible ( $0.11 \pm 0.02$  pmol/min/pmol P450 for EB and  $0.12 \pm 0.04$  pmol/min/pmol P450 for abamectin). Indeed, specific activities of CYP9A186wt were not significantly higher than the limits of detection (LOD) for EB or abamectin ( $0.10 \pm 0.01$  pmol/min/pmol P450 and  $0.13 \pm 0.02$  pmol/min/pmol P450, respectively) (Fig 5A). We observed no decrease of either EB or abamectin with microsomal protein fractions from cells infected with “non-insertion” control virus.

Two distinct metabolites were found in the CYP9A186-F116V microsomal reactions with EB (1.22 min and 1.37 min) and abamectin (1.65 min and 1.77 min) (S3 Fig). These metabolites were identified as (24' or 26') hydroxy-metabolites and O-desmethyl-metabolites by high resolution mass spectroscopy showing a quasi-molecular ion of m/z 902.5249 (hydroxy-EB, error = -1.3ppm, M + H), 872.5162 (O-desmethyl-EB, error = 0.8ppm, M + H), 906.5217 (hydroxy-abamectin, error = 0.8 ppm, M + NH<sub>4</sub>), 876.5096 (O-desmethyl-abamectin, error = -0.9 ppm, M+NH<sub>4</sub>) and a specific isotope peak ratio (n: n+1: n+2: n+3 = 10: 5: 1: 0.3) (S4 Fig).

These distinct hydroxy- and O-desmethyl-metabolites were only found in CYP9A186-F116V samples containing NADPH (S3A and S3D Fig), as the Ultra performance liquid chromatography-tandem mass spectrometer (UPLC-MS/MS) multireaction monitoring (MRM) signals corresponding to these two metabolites were not above the noise signal for the CYP9A186wt reaction samples (S3B and S3E Fig). These results confirm that unlike the CYP9A186-F116V samples, CYP9A186wt does not metabolize EB or abamectin under our experimental conditions.

## Discussion

The results reported here show the molecular genetic and biochemical basis of resistance to EB and abamectin in a laboratory-selected strain of beet armyworm from China. We first produced a chromosome-level *de novo* genome assembly and used BSA to identify the CYP9A gene cluster as the major locus associated with EB resistance. We then coupled CRISPR/Cas9 gene-editing to knockout regions within this gene cluster and mapped resistance to the CYP9A186 gene. Heterologous expression of recombinant CYP9A186 tested with *in vitro* metabolism assays confirmed that a natural mutation (T346G), resulting in the F116V substitution located in the substrate binding site 1 (SRS1) region of CYP9A186, causes resistance to both EB and abamectin. Hence, the combined use of CRISPR/Cas9 gene editing with BSA provided a rapid and highly efficient means to identify metabolic resistance genes responsible for avermectin insecticide resistance in the beet armyworm.

The evolution of pesticide resistance is an increasingly intractable problem affecting crop production worldwide [26]. To combat resistance, there is an urgent need for the rapid identification of not only markers for resistance, but also genes that cause pesticide resistance. Numerous studies point to target-site resistance and metabolic resistance as the two primary mechanisms by which insects evolve resistance to insecticides [27,28]. Target-site resistance involves alterations (e.g., mutations) in the insecticide target protein that reduce its sensitivity to insecticides [29]. For many insecticides, the primary insecticide targets are known, and identification of mutations can be directly assessed. In contrast, metabolic resistance can involve numerous different pathways consisting of unrelated gene products, making their identification difficult due to the high diversity of detoxification enzyme genes in insect pests [30,31].

One efficient way to overcome the above problem is to map resistance genes within finite regions of insect genomes by BSA [32]. To do this, a high-quality genome map is an absolute prerequisite for precise genetic localization. Hence, we generated a chromosome-level assembly of the beet armyworm genome by single-molecule real-time PacBio and Hi-C sequencing. Even with a precise genomic map, BSA can often be difficult to identify specific and functional genes responsible for the observed phenotype. Improvements in BSA will likely come as sequencing technologies and methods are further developed to increase their accuracy and shorten the confidence interval of genetic mapping, especially in regions of chromosomes dense with gene clusters and/or with low recombination rates. In addition to increased sequencing resolution, methods to improve how genetic loci identified by BSA are linked to resistance phenotypes via functional genomics are needed. Recently, the CRISPR/Cas9 genome editing system was shown to be effective in several lepidopterans, including *S. exigua* [33–35], providing a powerful functional genomics tool for the exploration of the complex biology of insecticide resistance. Here, we show for the first time that BSA used in conjunction with CRISPR/Cas9 gene editing allows for the specific and rapid identification of a gene responsible for metabolic insecticide resistance. Such information is important for tracking resistant beet armyworm individuals in the field and for making better informed decisions about the appropriate use of avermectin and related insecticides to delay the further spread of resistance. Furthermore, the identification of specific mutations involved in the structure and detoxification function of CYP9A186 may enable the rational design of new and more potent chemistries that target pest insects.

A P450 gene, CYP392A16, is highly over-expressed in the abamectin resistant Mar-ab strain of *T. urticae*, and this P450 can metabolize abamectin *in vitro* [17]. However, P450 point mutations that increase the metabolism of avermectins have not been previously reported. For cytochrome P450s, six substrate recognition sites (SRSs) are key to the acquisition of novel functions [36]. Site-directed mutagenesis studies previously showed that SRS1, SRS4, SRS5, and SRS6 possess amino acid residues that form the catalytic site, while SRS2 and SRS3 participate in the formation of the substrate access channel [37]. Conservative amino acid substitutions in these SRS regions can alter P450 activities and/or substrate profiles, and some substitutions in non-SRS regions on the proximal surface of a P450 can also improve kinetic properties due to alterations in interactions with the P450 redox partners [37]. Here, we show that the single amino acid substitution F116V in SRS1 of CYP9A186 is a gain-of-function mutation that enables enhanced metabolism of EB and abamectin. Although this Phe to Val substitution appears to be relatively minor, the replacement of the bulky phenyl aromatic ring for the isopropyl moiety apparently has a major effect on the CYP9A186 active site. Indeed, CYP9A186 belongs to the CYP3 clan of cytochrome P450s, which contains other well-characterized P450s, including the human CYP3A4 whose three-dimensional structure has been solved [38]. Alignment of CYP9A186 with CYP3A4 (Fig 4) shows that F116 in CYP9A186 likely corresponds with S119 of CYP3A4 within the B'-C loop. This residue likely protrudes into the active site, where it is in close proximity to the heme and opposite the C-terminal loop, allowing for direct interaction with substrate [38,39]. We speculate that the F116V substitution could reduce structural hindrance within the active site and allow for acceptance of a broader suite of substrates hydroxylated by the P450.

Changes in P450 transcription as well as structural changes appear to alter insecticide susceptibility and must be evaluated on a case-by-case basis. Recently, CYP321A8 was shown to contribute resistance to pyrethroid and organophosphate insecticides in *S. exigua* through overexpression regulated by cis- and trans-regulatory elements [40]. In *D. melanogaster*, transgenic overexpression of CYP6G2 leads to resistance to diazinon [41], whereas changes in transcript abundance was not involved for the closely related CYP6G1. It was rationalized that CYP6G2 has multiple hydrophobic substitutions that increased the size of its catalytic site

along with one positive-to-polar substitution that eliminated a potentially critical charge in SRS2 [37,42]. Similarly, structural changes in a P450 active site are important in *Anopheles gambiae*, where CYP6Z1 metabolizes DDT but CYP6Z2 does not [43,44]. Overlays of the predicted CYP6Z1 and CYP6Z2 structures indicated that R210 (SRS2), I298 and E302 (SRS4) of the CYP6Z2 protrude sufficiently enough into the CYP6Z2 catalytic site that it could no longer dock large hydrophobic substrates such as DDT [37]. In *Anopheles funestus*, alleles of *CYP6P9b* that confer high pyrethroid resistance are characterized by three point mutations, as well as constitutive overexpression [45]. More recently, Zimmer et al. [46] reported several point mutations were associated with gene duplications for CYP6ER1 gene variants of imidacloprid resistant *Nilaparvata lugens* strains. When two SRS5 mutations (A375del and A376G) were tested in transgenic *Drosophila*, imidacloprid resistance increased 20-fold. A similar level of resistance was obtained with the T318S (in SRS4) substitution. When this T318S substitution was combined with both A375del and A376P, this combination exhibited the highest resistance of all transgenic lines (35-fold). Such examples indicate that variation in both SRS regions and non-SRS regions of CYPs contribute to differences in metabolic capability and that gain-of-function mutations are common for several classes of insecticides.

*In vivo* and *in vitro* metabolism studies in vertebrates have identified three major oxidative metabolites of abamectin: 24-hydroxymethyl (24-OH), 26-hydroxymethyl (26-OH), and 3'-O-desmethyl abamectin [47]. In our study, CYP9A186-F116V also catalyzes the hydroxylation and O-demethylation of EB and abamectin, most likely also leading to the (24' or 26') hydroxy-metabolites and 3'-O-desmethyl-metabolites. In contrast, CYP392A16 from *T. urticae* did not O-demethylate abamectin, but rather catalyzed the hydroxylation of the insecticide to produce metabolites that were less toxic to *T. urticae* than the parental compound [17].

Alleles with fitness costs are known to have a decreased ability to survive and reproduce in the absence of the insecticide [48]. Previous research reported that the WH-EB strain of *S. exigua* developed 1,110-fold resistance to EB and that two or a few genes contributed to resistance [49]. Subsequently, the WH-EB strain was reared without further selection with EB. This strain now maintains a relatively stable resistance level (~300 fold) to EB. It is therefore likely that one or a few minor genes that contributed to resistance in the WH-EB strain were lost due to fitness costs, and only the mutated CYP9A186 with little or no associated fitness costs, was retained.

Here, our combined approach of coupling gene editing with BSA allowed for the rapid identification of *CYP9A186* as the primary gene involved in metabolic resistance to both EB and abamectin in a strain of the beet armyworm from China. The accuracy and speed at which we identified, mapped, and validated the functional role of *CYP9A186* suggests that such methods are highly adoptable for the rapid identification of resistance in many arthropod pest/insecticide systems. We can further use this information to proactively track beet armyworm avermectin resistance in field populations and make better informed decisions on the use of appropriate chemistries to further delay the evolution of insecticide resistance. Our findings also further enhance the understanding of avermectin resistance evolution and suggests that gain-of-function mutations in the SRS regions of P450 are important for the acquisition of novel detoxification functions. As such, future work to monitor *CYP9A186* for mutations in the field as well the design of specific compounds to target/inhibit *CYP9A186* could be important for resistance management of *S. exigua*.

## Materials and methods

### Insects

We used three strains of beet armyworm from China: WH-S, JZ-S, and WH-EB. WH-S is a avermectin-susceptible strain was obtained from Wuhan Institute of Vegetables (Hubei,

China) in 2009 and has been maintained in the laboratory for more than 20 years without exposure to any insecticides [6]. JZ-S is a susceptible strain obtained from the College of Plant Science and Technology, Huazhong Agriculture University (Wuhan, China) in 2018 and has been maintained in laboratory without exposure to any insecticides since its initial collection from Jingzhou, Hubei Province of China in 2003. WH-EB was derived from a field-collected strain with high levels of resistance to EB and is a near-isogenic strain introgressed with WH-S [49]. WH-EB was maintained without selection with EB since 2015.

### Genome sequencing

For genome sequencing, genomic DNA (gDNA) was extracted from two female pupae from the WH-S strain, including one pupa for both Illumina and PacBio whole genome sequencing (WGS) and the other pupa for Hi-C sequencing. gDNA was extracted using the Blood & Cell Culture DNA Mini Kit from Qiagen (Hilden, Germany). Long read libraries were constructed using a SMRTbell DNA Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA) from gDNA size-selected to approximately 15 kb and sequenced on a PacBio Sequel system with the Sequel Sequencing Kit 2.1 (Pacific Biosciences). Short read libraries corresponding to 350 bp insert size were prepared using the TruSeq DNA PCR-Free Low Throughput Library Preparation Kit (Illumina, San Diego, CA, USA). Libraries were subjected to 150-bp paired-end sequencing on the HiSeq X Ten platform (Illumina). Hi-C library preparation, including cross-linking the DNA, *Mbo* I restriction enzyme digestion, end repair, DNA cyclization, and DNA purification, was performed by Frasergen Co. Ltd. (Wuhan, China). All DNA extractions, library preparations, and DNA sequencing was performed by BGI (Shenzhen, China).

For RNA-Seq, total RNA was extracted using TRIzol Reagent (Thermo Fisher Scientific, Carlsbad, CA, USA) and sequencing libraries were constructed using the TruSeq RNA v2 Kit (Illumina). The library was paired-end sequenced using a 150 cycle high output sequencing kit, on an Illumina NextSeq 500 instrument.

### Genome assembly

Quality control was performed for the raw Illumina short reads using BBTools suite v38.49 [50]: duplicates were removed using `clumpy.sh`; low-quality bases were trimmed using `bbduk.sh`, with a minimum Phred quality score of 20 for both read sides, a minimum read/Ns length of 15/5 bp, a maximum length poly-A/G/C tail of 10 bp. Genome properties was surveyed using GenomeScope v1.0.0 [51] with 1,000 and 10,000 as the maximum *k*-mer coverage cutoffs.

Raw PacBio long reads were assembled into contigs using Flye v2.5 [52] and Falcon v1.3.0 [53]. Both assemblies were self-polished with a round of the Flye polishing module “—polish-target”. To improve contig contiguity, Flye and Falcon assemblies were merged with two rounds of quickmerge [54] following USAGE 2 (<https://github.com/mahulchak/quickmerge/wiki>; last accessed March 1, 2021). We removed redundant heterozygous contigs with Purge Haplotigs v1.1.0 [55] with the percent cutoff for identification of each contig as a haplotig set as 50 (-a 50) t. Non-redundant assembly was further polished with Illumina short reads using two rounds of Pilon v1.22 [56]; input BAM mapping files were generated using Minimap2 v2.17 [57] and SAMtools v1.9 [58].

To perform the chromosome-level concatenation, we aligned Hi-C read pairs to the genome, removed duplicates, and extracted the Hi-C contacts using Juicer v1.6.2 [59]. Pseudo-chromosomes were generated using 3D-DNA v180922 [60] in combination with a review step in Juicebox Assembly Tools module within Juicebox [59]. To identify sex chromosomes Z and W, we mapped Illumina short reads using BWA (DOI: [10.1093/bioinformatics/btp698](https://doi.org/10.1093/bioinformatics/btp698)), then

the average depth of coverage was calculated using SAMtools. Chromosomal nomenclature order and orientation were assigned by direct comparison with the *Bombyx mori* genome [61].

Sequences corresponding to contamination were removed by comparison with the nt and UniVec databases using HS-BLASTN v0.0.5 [62] and BLAST+ (i.e. blastn) v2.7.1 [63]. We assessed the assembly completeness using Benchmarking Universal Single-Copy Orthologs (BUSCO) [64] analyses against insect reference dataset (insect\_odb9, n = 1,658). We also mapped PacBio long reads and Illumina short reads to the final genome assembly with Minimap2.

### Genome annotation

To identify repetitive DNA elements, we constructed a *de novo* species-specific repeat library using RepeatModeler v1.0.11 [65], and then combined it with Dfam\_3.0 [66] and RepBase-20181026 databases [67] to generate a custom library. We then used RepeatMasker v4.0.9 [68] to mask repeats in the genome assembly with the combined custom library.

Protein-coding genes were predicted with the MAKER v2.31.10 pipeline [69] which can integrate *ab initio*, transcriptome- and protein homology-based evidence. For the *ab initio* gene predictions, we trained gene model parameters for Augustus v3.3 [70] and GeneMark-ET v4.38 [71] using BRAKER v2.1.0 [72], which incorporated RNA-seq data to accurately identify exon/intron boundaries. Input BAM alignment file for BRAKER was generated with HISAT2 v2.1.0 [73]. For the transcriptome-based evidence, a genome-guided method was applied to transcript assembly using StringTie v1.3.4 [74]. For the protein homology-based evidence, protein sequences of *Apis mellifera* (GCF\_003254395.2), *B. mori* (GCF\_000151625.1), *D. melanogaster* (GCF\_000001215.4), *Tribolium castaneum* (GCF\_000002335.3), *Helicoverpa armigera* (GCF\_002156985.1) and *Spodoptera litura* (GCF\_002706865.1) were downloaded directly from NCBI. Finally, repeat-masked genome and repeat annotations, assembled transcripts, protein sequences, and *ab initio* predicted gene models were passed to MAKER for automated gene structure predictions.

Gene functions were assigned to pro-coding gene models using Diamond v0.9.24 [75] against the UniProtKB (SwissProt + TrEMBL) database with the sensitive mode ‘—more-sensitive -e 1e-5’. We searched protein domains, Gene Ontology (GO) and Reactome pathway using InterProScan 5.41–78.0 [76] (-dp -f TSV, GFF3 -goterms -iprlookup -pa -t p -appl Pfam, Smart, Gene3D, Superfamily, CDD). We also annotated genes using eggNOG-mapper v2.0 [77] against the eggNOG v5.0 database [78] to obtain additional information of GO, Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology (KO), KEGG pathways, enzyme codes (ECs), and clusters of orthologous groups (COGs). Circos figure was produced by Tbtools v1.0 [79].

### Bulked segregant analysis and mapping by sequencing

To generate EB-resistant individuals for BSA, a single female moth from WH-EB was mated with a male moth from JZ-S to generate F<sub>1</sub> progeny. Because crossing over only occurs in males and EB resistance in WH-EB is incompletely dominant, a male F<sub>1</sub> progeny was then backcrossed with a JZ-S female to generate backcross progeny for selection on EB (Fig 2A). A total of 336 individuals (3<sup>rd</sup> instar larvae) from the backcross progeny were assayed on the diagnostic concentration of 0.01 µg/cm<sup>2</sup> EB, of which 134 individuals survived and comprised the BC-Sel family. We collected and pooled a single leg from each of the 120 BC-Sel survivors (e.g., pool for BSA consisted of a total of 120 moth legs) as well as the whole bodies for each of their parents. Total gDNA was extracted using a Blood & Cell Culture DNA Mini Kit (Qiagen). DNA was fragmented by sonication to obtain ~350 bp fragments and used for library

preparation using the NEBNext Ultra DNA Library Prep Kit (New England Biolabs, Ipswich, Massachusetts, USA). Libraries were sequenced on Illumina NovaSeq6000 PE-150bp sequencing system (Novogene Co. Ltd, Beijing, China). We sequenced 30x for each parent of the backcross family (male F<sub>1</sub> hybrid and female JZ-S), and 240x for the pooled 120 moth legs from the BC-Sel family (2x per individual).

The raw short-read data were filtered for quality by Trimmomatic software (version 0.36) with the following parameters “HEADCROP: 20 SLIDINGWINDOW: 4: 15 MINLEN: 50” [80] and mapping to the reference sequence using MEM module in Burrows-Wheeler Aligner software [81]. The resulting alignment files were transformed to binary files and sorted by SAMtools software [58]. PCR duplicates were identified by GATK [82] and indexed using SAMtools [58]. The HaplotypeCaller module in GATK was used to call variants [82]. SNP variants in each VCF files were filtered for quality using GATK with the following parameters [QD < 2.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0] [82].

Informative SNPs, which are heterozygous in the male F<sub>1</sub> hybrid and homozygous in the female parent (JZ-S), were selected to indicate the SNP segregation in the backcross progeny. All informative SNPs were expected to segregate 1:3 (for autosomes) or 1:2 (for Z chromosome) in the untreated backcross progeny. The absolute values of deviation between the detected frequencies in the BC-Sel pool and the expected frequencies (0.25 or 0.33) were calculated for each informative SNP. The mean frequency deviation values obtained from a sliding window analysis (1 Mb window with 100 Kb step) were plotted across all 31 chromosomes of the beet armyworm genome. Sliding windows with fewer than 80 informative SNPs were discarded in order to reduce noise. The mean frequency deviation values for the SNPs on Z chromosome were multiplied by a correction factor 0.75 (0.25/0.33) to make them comparable with that on autosomes.

### Guide RNA design for CRISPR/Cas9

Oligonucleotide primers used to make *CYP9A40*, *CYP9A107*, *CYP9A186* and *CYP9A98* single guide RNA (sgRNA) templates are shown in [S4 Table](#). Guide RNA was synthesized using the GeneArt Precision gRNA Synthesis kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer’s instructions. All PCR products were purified using a Qiagen PCR purification kit (Hilden, Germany) and used as templates for *in vitro* transcription. sgRNAs were purified and diluted in nuclease-free water to a concentration of ~2 μg μL<sup>-1</sup> and stored in aliquots at -80°C.

### Egg collection and microinjection

Freshly laid WH-EB eggs (within 2 h) were collected from gauze using 1% (v/v) sodium hypochlorite solution and rinsed with distilled water. Eggs were affixed onto microscope slides using double-sided adhesive tape [33,34]. Individual eggs were injected with approximately 2 nL of a solution containing two sgRNAs and the Cas9 protein (Thermo Fisher Scientific, Waltham, MA, USA) using a Nanoject III (Drummond, Broomall, PA, USA). The final concentration of each sgRNA was ~400 ng/μL and the concentration of Cas9 protein was 100 ng/μL. All microinjections were completed within 2 h (eggs less than 4 h old). Injected eggs were incubated at 25°C and 65% RH for 3–4 days until they hatched.

### CRISPR/Cas9 knockout of *CYP9A* genes

To detect CRISPR/Cas9-induced mutations, primers ([S4 Table](#)) were designed and used to amplify gDNA flanking the CRISPR target sites. We obtained gDNA from individuals using

the Multisource Genomic DNA Miniprep Kit (Axygen, New York, USA). PCR reactions contained 12.5  $\mu\text{L}$  of 2x EasyTaq mix (TransGen, Beijing, China), 1  $\mu\text{L}$  of each 10  $\mu\text{M}$  sense and antisense primer (S4 Table), 1  $\mu\text{L}$  gDNA, and 9.5  $\mu\text{L}$  ddH<sub>2</sub>O. PCR was performed at 94°C for 3 min, 32 cycles of (94°C 20 s, 56°C 20 s, 72°C 2 min) and 72°C for 7 min. Five  $\mu\text{L}$  of each PCR reaction was separated on 1.25% agarose gels stained with ethidium bromide. DNA bands were gel-purified, TA-cloned into the pClone007 blunt vector (Tsingke BioTech, Beijing, China), and Sanger sequenced by Tsingke BioTech.

To create strains homozygous for the Cas9-induced mutations, 30–80 pairs of G<sub>0</sub> moths were sib mated (S5 Table) and used to extract gDNA after laying eggs. PCR products flanking the two sgRNA target sites were amplified as indicated above to determine the genotypes and select lines to retain (S5–S7 Figs). At the 2<sup>nd</sup> instar larval stage, gDNA from 10 G<sub>1</sub> progeny was again extracted and used to detect the heritable mutations. Only progeny positive for the desired mutations were maintained (S5 Table). G<sub>2</sub> individuals that were heterozygous for the desired deletion mutation were mass crossed to produce G<sub>3</sub> (S8 Fig). Finally, homozygous G<sub>3</sub> individuals harboring the deletion mutation were then selected and mass crossed to establish each homozygous strain (S8 Fig).

### Sequencing *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98* cDNA

To assess mutations in RNA, we PCR amplified full-length cDNAs corresponding to *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98*. cDNA was prepared from total RNA extracted from pools of five 3<sup>rd</sup> instar larvae from WH-S and WH-EB. RNA was obtained using Trizol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instruction and treated with DNase I (TaKaRa, Shiga, Japan) to reduce genomic DNA contamination. First-strand cDNA synthesis was performed with 1  $\mu\text{g}$  of total RNA using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) cDNA synthesis kit (Promega, Madison, WI, USA).

Full-length cDNAs corresponding to *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98* were PCR amplified using specific primers (S6 Table) with reaction mixtures containing 10  $\mu\text{L}$  5x Q5 reaction buffer (New England Biolabs, Ipswich, MA, USA), 0.5  $\mu\text{L}$  dNTPs (10  $\mu\text{M}$ ), 1.25  $\mu\text{L}$  each of the sense and antisense primers (10  $\mu\text{M}$ ), 2  $\mu\text{L}$  cDNA, 0.25  $\mu\text{L}$  Q5 High-Fidelity DNA polymerase (New England Biolabs), and 9.75  $\mu\text{L}$  ddH<sub>2</sub>O in a final volume of 25  $\mu\text{L}$ . PCR conditions were 98°C 30 s, 35 cycles of 98°C 10 s, 60°C 20 s, 72°C for 2 min, and a final extension at 72°C for 2 min. PCR products were separated on 1.25% agarose gels in Tris-acetate-EDTA buffer and stained with ethidium bromide. PCR products of expected sizes were gel-purified with an AxyPrep Gel extraction kit (Axygen Biosciences, Union City, CA, USA), and then cloned into the pClone007 vector (Tsingke BioTech). *Escherichia coli* Trans1-T1 (Transgen Biotech) were transformed and inserts were sequenced by Tsingke BioTech. A total of 5 clones were sequenced for each cDNA for both WH-S and WH-EB.

### Quantitative real-time PCR analysis of *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98*

For RT-qPCR analysis of *CYP9A107*, *CYP9A27*, *CYP9A11*, *CYP9A186* and *CYP9A98*, we isolated total RNA from the midgut and fat bodies of individual WH-S and WH-EB 5<sup>th</sup> instars. First-strand cDNA was synthesized using 2  $\mu\text{g}$  of total RNA using PrimeScript Reverse Transcriptase (TaKaRa, Dalian, China) according to the manufacturer's protocol. The reaction cycling conditions for the five *CYP9As* are identical to those used by Shi et al. [83]. Four biological replicates were done for each gene. qPCR results were analyzed using the  $2^{-\Delta\Delta\text{Ct}}$



method [84]. Data were normalized to the geometric mean of two housekeeping genes (*β-actin* and *GADPH*). Significant differences in transcript abundance were determined using the Student's *t*-test ( $\alpha = 0.05$ ) using SPSS 20.0 software (SPSS Inc., Chicago, IL, USA).

### Chemicals and bioassays

Emamectin Benzoate (EB) (~93.4%B1a, CAS: 155569-91-8, 98%), 7-benzyloxy-4-trifluoromethyl coumarin (BFC) and reduced nicotinamide adenine dinucleotide phosphate (NADPH) were purchased from Sigma-Aldrich company (St. Louis, Missouri, USA). Abamectin (B1a, CAS: 65195-55-3, 97%) was purchased from Toronto Research Chemicals (North York, Canada). Solvents for LC-MS (ammonium acetate, formic acid and acetonitrile) were obtained from Fisher Scientific (Pittsburgh, Pennsylvania, USA).

A series of gradient concentrations of insecticides diluted from stock solution with 0.1% (w/v) Triton X-100 water were prepared. In each well of the 24-well plate, a liquid artificial diet (~120  $\mu$ L) was supplied. One hundred microliters of the insecticide solution was added to the dietary surface of every single well after the diet cooled and solidified. A third-instar larva was placed in each well. For each concentration, 24 larvae were treated. Triton X-100 (0.1%, w/v) was employed to process the control group. Two days later, larvae were checked and considered dead if they did not move after gentle prodding. The data were analyzed using DPS 7.05 software (Zhejiang University, Hangzhou, China). When the 95% fiduciary limits of the  $LC_{50}$  values did not overlap, the  $LC_{50}$  values were considered to vary significantly.

### Heterologous expression of CYP9A186

Full-length cDNAs corresponding to the ORFs of wild-type *CYP9A186* (GenBank accession no. MN179467) and mutant *CYP9A186* (*CYP9A186-F116V*) (GenBank accession no. MN179472) were subcloned into pFastBac1 (Invitrogen) from WH-S and WH-EB, respectively. Controls included the beet armyworm cytochrome P450 reductase (SeCPR, GenBank accession no. MN179471) subcloned from WH-S and cells infected with virus containing empty pFastBac1 vector (non-insertion negative control). Heterologous protein expression was carried out using the Bac-to-Bac system (Invitrogen) as previously described [83]. Briefly, Hi5 cells were maintained at 27°C with 130 rpm mixing in Express Five serum-free medium (Invitrogen). P3 virus was added when Hi5 cell density reached  $2-3 \times 10^6$  cells  $mL^{-1}$ . Cells were co-infected with recombinant virus (either *CYP9A186*, *CYP9A186-F116V*, or non-insertion control) and SeCPR with multiplicity of infection (MOI) of 2 and 0.2, respectively. Ferric citrate (0.1 mM) and 5'-aminolevulinic acid hydrochloride (0.1 mM) were added to the medium along with virus and 24 h post infection. Microsomes were prepared by differential centrifugation. Total microsomal protein content was determined using the Bradford method [85] and the amount of recombinant P450 was quantified using the reduced CO-difference spectral assay [86].

### *In vitro* metabolism of model substrates and insecticides

To observe the metabolism exhibited by *CYP9A186* and *CYP9A186-F116V*, *in vitro* enzyme assays were performed on microsomal fractions containing recombinant P450/SeCPR (or non-insertion control/SeCPR). All assays were performed in 0.1 M potassium phosphate buffer (pH 7.4) with various substrates in a total reaction volume of 200  $\mu$ L.

To test for the presence of active recombinant cytochrome P450s (indicating correct protein folding), 25 mM 7-benzyloxy-4-trifluoromethyl coumarin (BFC) (dissolved in DMSO) was added to 20 pmol aliquots of each microsomal protein sample in black, flat-bottom 96-well plates for 5 min at 30°C. Reactions were initiated by adding 10  $\mu$ L of 10 mM NADPH.

The formation of metabolized product was monitored using a Spectra Max M5 plate reader (Molecular Devices) as previously described [83]. Controls included samples without NADPH and microsomal samples from “no insert”-infected cells (non-insertion control). Specific activities were corrected by subtracting the background (activity of non-insertion control) and expressed as pmol product per minute per pmol P450.

To measure the metabolism of EB and abamectin directly, 10 pmol aliquots of each microsomal protein sample was incubated with an NADPH regenerating system (1.3 mM NADP<sup>+</sup>, 3.3 mM glucose-6-phosphate, 3.3 mM MgCl<sub>2</sub> and 0.4 U mL<sup>-1</sup> glucose-6-phosphate dehydrogenase) and 10 μM of either insecticide (2 μL dissolved in DMSO). Reactions were pre-incubated for 5 min at 30°C before adding insecticide substrate. Reactions were carried out in an orbital shaking incubator at 30°C at 1,200 rpm. Reactions were stopped by adding 200 μL acetonitrile and 600 μL dilution buffer (50% potassium phosphate buffer and 50% acetonitrile). Each reaction was further incubated/shaken for an additional 20 min. Samples were centrifuged at 18,000 × g for 10 min and supernatants (800 μL) were transferred to HPLC vials and analyzed immediately. Control samples (e.g., without NADPH and non-insertion negative controls) containing equal amount of total protein were subsequently tested. The appearance of degraded insecticides indicates metabolic enzymatic activity. Specific activities were corrected by subtracting the background (activity of non-insertion negative control) and expressed as pmol substrate/min/pmol P450.

### Identification of insecticide metabolites by UPLC-MS/MS

The degradation of EB and abamectin was monitored and quantified by UPLC-MS/MS. Samples (0.5 μL each) were separated by Acquity BEH C8 column (2.1 × 50 mm, 1.7 μm particle size, Waters, Milford, Massachusetts) using a Waters Acquity UPLC system (Waters ACQUITY UPLC I-Class) and eluted with a gradient of mobile phase consisting of A: 10 mM ammonium acetate [+0.1% (v/v) formic acid] and B: acetonitrile, with a constant flow rate of 0.3 mL min<sup>-1</sup>. The gradient elution conditions were as follows: 0 min A: B 50: 50; 0.3 min A: B 50: 50; 2 min A: B 5:95; 2.5 min A: B 5: 95; 2.6 min A: B 0:100; 3 min A: B 0: 100; 3.1 min acetonitrile: Water 50: 50; 5 min acetonitrile: Water 50: 50. Samples separated by UPLC were directly analyzed using a tandem triple quadrupole mass-spectrometer (Waters Xevo TQ-S micro, Waters, Milford, Massachusetts) and run in positive ESI mode with multireaction monitoring (MRM) (S7 Table).

For identification of the EB and abamectin metabolites, samples were separated using a Shimadzu LC system (LC20ADXR with Agilent Poroshell 120 EC-C18 column 2.1 mm × 50 mm, 2.7 μm) and the same mobile phases and gradient as outlined above. The partitioned samples were then analyzed by high-resolution mass spectroscopy (Triple ToF 5600+, AB Sciex, USA) under +ESI mode with collision energy at 50V.

### Supporting information

**S1 Fig. Schematic diagram of the sgRNA targeting sites and the CRISPR/Cas9 CYP9A186 mutagenesis.** (A) Schematic diagram of the sgRNA-targeting sites. The gray line indicates the genome locus of *CYP9A186* and the boxes represent the exons of *CYP9A186*. The sgRNA-targeting site was located on the sense strand of exon 1. The sgRNA-targeting sequence is shown in purple, and the protospacer adjacent motif (PAM) sequence is in red. (B) Representative chromatograms of PCR-product sequencing in G<sub>1</sub> individuals and the *CYP9A186* knockout strain (dA186) showing presence of 4-bp indel mutation. (C) Diagram detailing crossing scheme used to obtain the homozygous *CYP9A186* knockout strain (+, wild-type; -, mutant). (TIFF)

**S2 Fig. Reduced CO difference spectra of recombinant CYP9A186 F116V mutant, wild type and non-insertion control.**

(TIFF)

**S3 Fig. Formation of hydroxy- and O-desmethy-metabolites in *in vitro* metabolism (MRM spectra).** Metabolism of emamectin benzoate: **A**, samples with CYP9A186-F116V; **B**, samples with CYP9A186wt; **C**, samples with non-insertion control (CK). Metabolism of abamectin: **D**, samples with CYP9A186-F116V; **E**, samples with CYP9A186wt; **F**, samples with non-insertion control (CK). MRM spectra of parent compound (emamectin benzoate and abamectin), hydroxy-metabolite (-OH) and O-desmethyl-metabolite (-O-CH<sub>3</sub>) were shown in the top, middle, and bottom respectively within each block. Red arrows indicate metabolites.

(TIFF)

**S4 Fig. ESI-TOF high resolution MS spectrum of metabolites resulting from incubations of emamectin benzoate and abamectin with recombinantly expressed CYP9A186-F116V.**

**(A)** Mass spectra of the metabolite hydroxyl-emamectin benzoate (HO-groups correspond to either 24' or 26' position). **(B)** Mass spectra of the metabolite O-desmethyl-emamectin benzoate. **(C)** Mass spectra of the metabolite hydroxyl-abamectin (HO-groups correspond to either 24' or 26' position). **(D)** Mass spectra of the metabolite O-desmethyl-abamectin.

(TIFF)

**S5 Fig. CRISPR/Cas9-based knockout of the CYP9A cluster (dA40-A98) in the beet armyworm WH-EB strain.**

**(A)** sgRNA targeting site of *CYP9A40* and *CYP9A98* genes and the two primer pairs for allele-specific PCR detection. Target sequences and protospacer adjacent motifs (PAMs) are shown in light blue and in red, respectively. The positions of the two sgRNAs (sgRNA-A40 and sgRNA-A98) and a representative chromatogram of direct sequencing of PCR products of individuals from the dA40-A98 strain with the primer pair A40F/A98R are shown. **(B)** Genotyping of individual *S. exigua* for deletion of the *CYP9A* cluster according to banding patterns of the PCR products amplified with a set of five primer pairs. M, 2000 bp MW Marker; Lane 1, A40F/A98R; Lane 2, A40F/A40R; Lane 3, A9F/A9R; Lane 4, A186F1/A186R1; and Lane 5, A98F/A98R.

(TIFF)

**S6 Fig. CRISPR/Cas9-based knockout of the CYP9A cluster (dA40-A107) in the beet armyworm WH-EB strain.**

**(A)** sgRNA targeting site of *CYP9A40* and *CYP9A107* genes and the two primer pairs for allele-specific PCR detection. Target sequences and protospacer adjacent motifs (PAMs) are shown in light blue and in red, respectively. The positions of the two sgRNAs (sgRNA-A40 and sgRNA-A107) and a representative chromatogram of direct sequencing of PCR products of individuals from the dA40-A107 strain with the primer pair A40F/A107R are shown. **(B)** Genotyping of individual *S. exigua* for deletion from *CYP9A40* to *CYP9A107* according to banding patterns of the PCR products amplified with a set of five primer pairs. M, 2000 bp MW Marker; Lane 1, A40F/A107R; Lane 2, A40F/A40R; Lane 3, A107F/A107R; Lane 4, A9F/A9R; and Lane 5, A186F2/A186R2.

(TIFF)

**S7 Fig. CRISPR/Cas9-based knockout of the CYP9A cluster (dA107-A98) in the beet armyworm WH-EB strain.**

**(A)** sgRNA targeting site of *CYP9A107* and *CYP9A98* genes and the two primer pairs for allele-specific PCR detection. Target sequences and protospacer adjacent motifs (PAMs) are shown in light blue and in red, respectively. The positions of the two sgRNAs (sgRNA-A107 and sgRNA-A98) and a representative chromatogram of direct

sequencing of PCR products of individuals from the dA107-A98 strain with the primer pair A107F/A98R are shown. (B) Genotyping of individual *S. exigua* for deletion from *CYP9A107* to *CYP9A98* according to banding patterns of the PCR products amplified with a set of five primer pairs. M, 2000 bp MW Marker; Lane 1, A107F/A98R; Lane 2, A107F/A107R; Lane 3, A186F2/A186R2; and Lane 4, A98F/A98R.

(TIFF)

**S8 Fig. Diagram detailing crossing scheme used to obtain homozygous knockout strains of *Spodoptera exigua*.**

(TIFF)

**S1 Table. Genome survey for *Spodoptera exigua*.**

(DOCX)

**S2 Table. Summary of each assembly version for *Spodoptera exigua*.**

(DOCX)

**S3 Table. Repeat annotation in *Spodoptera exigua*.**

(DOCX)

**S4 Table. Primers for template DNA for in vitro transcription of sgRNAs and amplifying genomic DNA fragment of CYP9A genes.**

(DOCX)

**S5 Table. Mutagenesis induced by gRNA/Cas9.**

(DOCX)

**S6 Table. Primers for amplifying *SeCYP9A* subfamily genes and analysis *SeCYP9A* subfamily genes expression.**

(DOCX)

**S7 Table. Analysis parameters for substrates and metabolites detection.**

(DOCX)

## Author Contributions

**Conceptualization:** Yihua Yang, Yidong Wu.

**Data curation:** Yayun Zuo, Yu Shi, Feng Zhang, Fang Guan, Jianpeng Zhang, René Feyereisen, Jeffrey A. Fabrick, Yidong Wu.

**Funding acquisition:** Yihua Yang.

**Investigation:** Yayun Zuo, Yu Shi, Feng Zhang, Fang Guan, Jianpeng Zhang.

**Resources:** Yihua Yang.

**Supervision:** Yihua Yang, Yidong Wu.

**Writing – original draft:** Yayun Zuo, Yu Shi, Feng Zhang.

**Writing – review & editing:** René Feyereisen, Jeffrey A. Fabrick, Yidong Wu.

## References

1. Wolstenholme AJ. Glutamate-gated chloride channels. *Journal of Biological Chemistry*. 2012; 287: 40232–40238. <https://doi.org/10.1074/jbc.R112.406280> PMID: [23038250](https://pubmed.ncbi.nlm.nih.gov/23038250/)

2. Ishaaya I, Kontsedalov S, Horowitz AR. Emamectin, a novel insecticide for controlling field crop pests. *Pest Management Science*. 2002; 58: 1091–1095. <https://doi.org/10.1002/ps.535> PMID: 12449526
3. Kim JH, Gellatly KJ, Lueke B, Kohler M, Nauen R, Murenzi E, et al. Detoxification of ivermectin by ATP binding cassette transporter C4 and cytochrome P450 monooxygenase 6CJ1 in the human body louse, *Pediculus humanus humanus*. *Insect Molecular Biology*. 2018; 27: 73–82. <https://doi.org/10.1111/imb.12348> PMID: 28960749
4. Artho R, Schnyder M, Kohler L, Torgerson PR, Hertzberg H. Avermectin-resistance in gastrointestinal nematodes of Boer goats and Dorper sheep in Switzerland. *Veterinary Parasitology*. 2007; 144: 68–73. <https://doi.org/10.1016/j.vetpar.2006.09.032> PMID: 17088023
5. Dermauw W, Ilias A, Riga M, Tsagkarakou A, Grbic M, Tirry L, et al. The cys-loop ligand-gated ion channel gene family of *Tetranychus urticae*: implications for acaricide toxicology and a novel mutation associated with abamectin resistance. *Insect Biochemistry and Molecular Biology*. 2012; 42: 455–465. <https://doi.org/10.1016/j.ibmb.2012.03.002> PMID: 22465149
6. Che W, Shi T, Wu YD, Yang YH. Insecticide resistance status of field populations of *Spodoptera exigua* (Lepidoptera: Noctuidae) from China. *Journal of Economic Entomology*. 2013; 106: 1855–1862. <https://doi.org/10.1603/ec13128> PMID: 24020303
7. Lebedev G, Abo-Moch F, Gafni G, Ben-Yakir D, Ghanim M. High-level of resistance to spinosad, emamectin benzoate and carbosulfan in populations of *Thrips tabaci* collected in Israel. *Pest Management Science*. 2013; 69: 274–277. <https://doi.org/10.1002/ps.3385> PMID: 22926932
8. Ghosh R, Andersen EC, Shapiro JA, Gerke JP, Kruglyak L. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science*. 2012; 335: 574–578. <https://doi.org/10.1126/science.1214318> PMID: 22301316
9. Kane NS, Hirschberg B, Qian S, Hunt D, Thomas B, Brochu R, et al. Drug-resistant *Drosophila* indicate glutamate-gated chloride channels are targets for the antiparasitics nodulisporic acid and ivermectin. *Proceedings of the National Academy of Sciences of the United States of America*. 2000; 97: 13949–13954. <https://doi.org/10.1073/pnas.240464697> PMID: 11095718
10. Wang XL, Wang R, Yang YH, Wu SW, O'Reilly AO, Wu YD. A point mutation in the glutamate-gated chloride channel of *Plutella xylostella* is associated with resistance to abamectin. *Insect Molecular Biology*. 2016; 25: 116–125. <https://doi.org/10.1111/imb.12204> PMID: 26592158
11. Wang XL, Puinean AM, Williamson MS, Smelt CL, Millar NS, Wu YD. Mutations on M3 helix of *Plutella xylostella* glutamate-gated chloride channel confer unequal resistance to abamectin by two different mechanisms. *Insect Biochemistry and Molecular Biology*. 2017; 86: 50–57. <https://doi.org/10.1016/j.ibmb.2017.05.006> PMID: 28576654
12. Kwon DH, Yoon KS, Clark JM, Lee SH. A point mutation in a glutamate-gated chloride channel confers abamectin resistance in the two-spotted spider mite, *Tetranychus urticae* Koch. *Insect Molecular Biology*. 2010; 19: 583–591. <https://doi.org/10.1111/j.1365-2583.2010.01017.x> PMID: 20522121
13. Yoon KS, Neison JO, Clark JM. Selective induction of abamectin metabolism by dexamethasone, 3-methylcholanthrene, and phenobarbital in Colorado potato beetle, *Leptinotarsa decemlineata*. *Pesticide Biochemistry and Physiology*. 2002; 73: 74–86.
14. Wang LH, Wu, YD. Cross-resistance and biochemical mechanisms of abamectin resistance in the B-type *Bemisia tabaci*. *Journal of Applied Entomology*. 2007; 131: 98–103.
15. Qian L, Cao G, Song J, Yin Q, Han Z. Biochemical mechanisms conferring cross-resistance between tebufenozide and abamectin in *Plutella xylostella*. *Pesticide Biochemistry and Physiology*. 2008; 91: 175–179.
16. Stumpf N, Nauen R. Biochemical markers linked to abamectin resistance in *Tetranychus urticae* (Acari: Tetranychidae). *Pesticide Biochemistry and Physiology*. 2002; 72: 111–121.
17. Riga M, Tsakireli D, Ilias A, Morou E, Myridakis A, Stephanou EG, et al. Abamectin is metabolized by CYP392A16, a cytochrome P450 associated with high levels of acaricide resistance in *Tetranychus urticae*. *Insect Biochemistry and Molecular Biology*. 2014; 46: 43–53. <https://doi.org/10.1016/j.ibmb.2014.01.006> PMID: 24463358
18. Konanz S, Nauen R. Purification and partial characterization of a glutathione S-transferase from the two-spotted spider mite, *Tetranychus urticae*. *Pesticide Biochemistry and Physiology*. 2004; 79: 49–57.
19. Ishtiaq M, Saleem MA, Razaq M. Monitoring of resistance in *Spodoptera exigua* (Lepidoptera: Noctuidae) from four districts of the Southern Punjab, Pakistan to four conventional and six new chemistry insecticides. *Crop Protection*. 2012; 33: 13–20.
20. Becker A, Chao DY, Zhang X, Salt DE, Baxter I. Bulk segregant analysis using single nucleotide polymorphism microarrays. *PLoS One*. 2011; 6: e15993. <https://doi.org/10.1371/journal.pone.0015993> PMID: 21297997

21. Traut W, Sahara K, Marec F. Sex chromosomes and sex determination in Lepidoptera. *Sexual Development*. 2007; 1: 332–346. <https://doi.org/10.1159/000111765> PMID: [18391545](#)
22. Cheng TC, Wu JQ, Wu YQ, Chilukuri RV, Huang LH, Yamamoto K, et al. Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. *Nature Ecology & Evolution*. 2017; 1: 1747–1756.
23. Giraudo M, Hilliou F, Fricaux T, Audant P, Feyereisen R, Le Goff G. Cytochrome P450s from the fall armyworm (*Spodoptera frugiperda*): responses to plant allelochemicals and pesticides. *Insect Molecular Biology*. 2015; 24: 115–28. <https://doi.org/10.1111/imb.12140> PMID: [25315858](#)
24. Hu B, Zhang SH, Ren MM, Tian XR, Wei Q, Mburu DK, et al. The expression of *Spodoptera exigua* P450 and UGT genes: tissue specificity and response to insecticides. *Insect Science*. 2019; 26: 199–216. <https://doi.org/10.1111/1744-7917.12538> PMID: [28881445](#)
25. Gimenez S, Abdelgaffar H, Goff GL, Hilliou F, Blanco CA, Hanniger S, et al. Adaptation by copy number variation increases insecticide resistance in the fall armyworm. *Communications Biology*. 2020; 3: 664. <https://doi.org/10.1038/s42003-020-01382-6> PMID: [33184418](#)
26. Gould F, Brown ZS, Kuzma J. Wicked evolution: Can we address the sociobiological dilemma of pesticide resistance? *Science*. 2018; 360: 728–732. <https://doi.org/10.1126/science.aar3780> PMID: [29773742](#)
27. 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. <https://doi.org/10.1016/j.pestbp.2015.01.004> PMID: [26047113](#)
28. Liu N. Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. *Annual Review of Entomology*. 2015; 60: 537–559. <https://doi.org/10.1146/annurev-ento-010814-020828> PMID: [25564745](#)
29. Zimmer CT, Garrood WT, Puinean AM, Eckel-Zimmer M, Williamson MS, Davies TE, et al. A CRISPR/Cas9 mediated point mutation in the alpha 6 subunit of the nicotinic acetylcholine receptor confers resistance to spinosad in *Drosophila melanogaster*. *Insect Biochemistry and Molecular Biology*. 2016; 73: 62–69. <https://doi.org/10.1016/j.ibmb.2016.04.007> PMID: [27117524](#)
30. Oakeshott JG, Farnsworth CA, East PD, Scott C, Han Y, Wu Y, et al. How many genetic options for evolving insecticide resistance in *Heliothine* and *Spodopteran* pests? *Pest management science*. 2013; 69: 889–896. <https://doi.org/10.1002/ps.3542> PMID: [23526801](#)
31. Wang HD, Shi Y, Wang L, Liu S, Wu SW, Yang YH, et al. CYP6AE gene cluster knockout in *Helicoverpa armigera* reveals role in detoxification of phytochemicals and insecticides. *Nature Communications*. 2018; 9: 4820. <https://doi.org/10.1038/s41467-018-07226-6> PMID: [30446639](#)
32. Kurlovs AH, Snoeck S, Kosterlitz O, Van Leeuwen T, Clark RM. Trait mapping in diverse arthropods by bulked segregant analysis. *Current Opinion in Insect Science*. 2019; 36: 57–65. <https://doi.org/10.1016/j.cois.2019.08.004> PMID: [31499416](#)
33. Zuo YY, Huang JL, Wang J, Feng Y, Han TT, Wu YD, et al. Knockout of a P-glycoprotein gene increases susceptibility to abamectin and emamectin benzoate in *Spodoptera exigua*. *Insect Molecular Biology*. 2018; 27: 36–45. <https://doi.org/10.1111/imb.12338> PMID: [28753233](#)
34. Zuo YY, Wang H, Xu YJ, Huang JL, Wu SW, Wu YD, et al. CRISPR/Cas9 mediated G4946E substitution in the ryanodine receptor of *Spodoptera exigua* confers high levels of resistance to diamide insecticides. *Insect Biochemistry and Molecular Biology*. 2017; 89: 79–85. <https://doi.org/10.1016/j.ibmb.2017.09.005> PMID: [28912111](#)
35. Huang JL, Xu YJ, Zuo YY, Yang YH, Tabashnik BE, Wu YD. Evaluation of five candidate receptors for three Bt toxins in the beet armyworm using CRISPR-mediated gene knockouts. *Insect Biochemistry and Molecular Biology*. 2020; 121: 103361. <https://doi.org/10.1016/j.ibmb.2020.103361> PMID: [32199887](#)
36. Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *Journal of Biological Chemistry*. 1992; 267: 83–90. PMID: [1730627](#)
37. Schuler MA, Berenbaum MR. Structure and function of cytochrome P450S in insect adaptation to natural and synthetic toxins: insights gained from molecular modeling. *Journal of Chemical Ecology*. 2013; 39: 1232–1245. <https://doi.org/10.1007/s10886-013-0335-7> PMID: [24036972](#)
38. Park H, Lee S, Sub J. Structural and dynamical basis of broad substrate specificity, catalytic mechanism, and inhibition of cytochrome P450 3A4. *Journal of the American Chemical Society*. 2005; 127: 13634–13642. <https://doi.org/10.1021/ja053809q> PMID: [16190729](#)
39. Sevrioukova IF, Thomas L, Poulos TL. Structural basis for regiospecific midazolam oxidation by human cytochrome P450 3A4. *Proceedings of the National Academy of Sciences of the United States of America*. 2017; 114: 486–491. <https://doi.org/10.1073/pnas.1616198114> PMID: [28031486](#)

40. Hu B, Huang H, Hu S, Ren M, Wei Q, Tian X, et al. Changes in both trans- and cis-regulatory elements mediate insecticide resistance in a lepidopteron pest, *Spodoptera exigua*. *PLoS Genetics*. 2021; 17: e1009403. <https://doi.org/10.1371/journal.pgen.1009403> PMID: 33690635
41. Daborn PJ, Lumb C, Boey A, Wong W, Ffrench-Constant RH, Batterham P. Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochemistry and Molecular Biology*. 2007; 37: 512–519. <https://doi.org/10.1016/j.ibmb.2007.02.008> PMID: 17456446
42. Jones RT, Bakker SE, Stone D, Shuttleworth SN, Boundy S, McCart C, et al. Homology modeling of *Drosophila* cytochrome P450 enzymes associated with insecticide resistance. *Pest Management Science*. 2010; 66: 1106–1115. <https://doi.org/10.1002/ps.1986> PMID: 20583201
43. Chiu TL, Wen Z, Rupasinghe SG, Schuler MA. Comparative molecular modeling of *Anopheles gambiae* CYP6Z1, a mosquito P450 capable of metabolizing DDT. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105: 8855–8860. <https://doi.org/10.1073/pnas.0709249105> PMID: 18577597
44. McLaughlin LA, Niaz U, Bibby J, David JP, Vontas J, Hemingway J, et al. Characterization of inhibitors and substrates of *Anopheles gambiae* CYP6Z2. *Insect Molecular Biology*. 2008; 17: 125–135. <https://doi.org/10.1111/j.1365-2583.2007.00788.x> PMID: 18353102
45. Ibrahim SS, Riveron JM, Bibby J, Irving H, Yunta C, Paine MJ, et al. Allelic variation of cytochrome P450s drives resistance to bednet insecticides in a major malaria vector. *PLoS Genetics*. 2015; 11: e1005618. <https://doi.org/10.1371/journal.pgen.1005618> PMID: 26517127
46. Zimmer CT, Garrod WT, Singh KS, Randall E, Lueke B, Gutbrod O, et al. Neofunctionalization of duplicated P450 genes drives the evolution of insecticide resistance in the brown planthopper. *Current Biology*. 2018; 28: 268–274. <https://doi.org/10.1016/j.cub.2017.11.060> PMID: 29337073
47. Zeng Z, Andrew NW, Woda JM, Halley BA, Crouch LS, Wang RW. Role of cytochrome P450 isoforms in the metabolism of abamectin and ivermectin in rats. *Journal of Agricultural and Food Chemistry*. 1996; 44: 3374–3378.
48. Gassmann AJ, Carrière Y, Tabashnik BE. Fitness costs of insect resistance to *Bacillus thuringiensis*. *Annual review of entomology*. 2009; 54: 147–163. <https://doi.org/10.1146/annurev.ento.54.110807.090518> PMID: 19067630
49. Che W, Huang JL, Guan F, Wu YD, Yang YH. Cross-resistance and inheritance of resistance to emamectin benzoate in *Spodoptera exigua* (Lepidoptera: Noctuidae). *Journal of Economic Entomology*. 2015; 108: 2015–2020. <https://doi.org/10.1093/jee/tov168> PMID: 26470348
50. Bushnell B. BBtools. 2014; Retrieved from <https://sourceforge.net/projects/bbmap/>
51. Vurture GW, Sedlazeck FJ, Nattestad M, Underwood CJ, Fang H, Gurtowski J, Schatz MC. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*. 2017; 33: 2202–2204. <https://doi.org/10.1093/bioinformatics/btx153> PMID: 28369201
52. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*. 2019; 37: 540–546 <https://doi.org/10.1038/s41587-019-0072-8> PMID: 30936562
53. Chin CS, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, et al. Phased diploid genome assembly with single-molecule real-time sequencing. *Nature Methods*. 2016; 13: 1050–1054. <https://doi.org/10.1038/nmeth.4035> PMID: 27749838
54. Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. Contiguous and accurate *de novo* assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Research*. 2016; 44: e147. <https://doi.org/10.1093/nar/gkw654> PMID: 27458204
55. Roach MJ, Schmidt SA, Borneman AR. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*. 2018; 19: 460. <https://doi.org/10.1186/s12859-018-2485-7> PMID: 30497373
56. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*. 2014; 9: e112963. <https://doi.org/10.1371/journal.pone.0112963> PMID: 25409509
57. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018; 34: 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191> PMID: 29750242
58. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*. 2009; 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
59. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems*. 2016; 3: 95–98. <https://doi.org/10.1016/j.cels.2016.07.002> PMID: 27467249

60. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, et al. *De novo* assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*. 2017; 356: 92–95. <https://doi.org/10.1126/science.aal3327> PMID: 28336562
61. Kawamoto M, Jouraku A, Toyoda A, Yokoi K, Minakuchi Y, Katsuma S, et al. High-quality genome assembly of the silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*. 2019; 107: 53–62. <https://doi.org/10.1016/j.ibmb.2019.02.002> PMID: 30802494
62. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Research*. 2015; 43: 7762–7768. <https://doi.org/10.1093/nar/gkv784> PMID: 26250111
63. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009; 10: 421. <https://doi.org/10.1186/1471-2105-10-421> PMID: 20003500
64. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*. 2018; 35: 543–548. <https://doi.org/10.1093/molbev/msx319> PMID: 29220515
65. Smit AFA, Hubley R. RepeatModeler Open-1.0. 2008–2015; Retrieved from <http://www.repeatmasker.org>
66. Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, et al. The Dfam database of repetitive DNA families. *Nucleic Acids Research*. 2016; 44: D81–D89. <https://doi.org/10.1093/nar/gkv1272> PMID: 26612867
67. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*. 2015; 6: 11. <https://doi.org/10.1186/s13100-015-0041-9> PMID: 26045719
68. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013–2015; Retrieved from <http://www.repeatmasker.org>
69. Holt C, Yandell M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*. 2011; 12: 491. <https://doi.org/10.1186/1471-2105-12-491> PMID: 22192575
70. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research*. 2004; 32: W309–W312. <https://doi.org/10.1093/nar/gkh379> PMID: 15215400
71. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*. 2005; 33: 6494–6506. <https://doi.org/10.1093/nar/gki937> PMID: 16314312
72. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*. 2016; 32: 767–769. <https://doi.org/10.1093/bioinformatics/btv661> PMID: 26559507
73. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*. 2015; 12: 357–360. <https://doi.org/10.1038/nmeth.3317> PMID: 25751142
74. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*. 2015; 33: 290–295. <https://doi.org/10.1038/nbt.3122> PMID: 25690850
75. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*. 2015; 12: 59–60. <https://doi.org/10.1038/nmeth.3176> PMID: 25402007
76. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, et al. InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Research*. 2017; 45: D190–D199. <https://doi.org/10.1093/nar/gkw1107> PMID: 27899635
77. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Molecular Biology and Evolution*. 2017; 34: 2115–2122. <https://doi.org/10.1093/molbev/msx148> PMID: 28460117
78. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*. 2017; 45: D309–D314.
79. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*. 2020; 13: 1194–1202. <https://doi.org/10.1016/j.molp.2020.06.009> PMID: 32585190
80. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
81. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013; 1303.3997.



82. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. 2010; 20: 1297–1303. <https://doi.org/10.1101/gr.107524.110> PMID: [20644199](https://pubmed.ncbi.nlm.nih.gov/20644199/)
83. Shi Y, Wang HD, Liu Z, Wu SW, Yang YH, Feyereisen R, et al. Phylogenetic and functional characterization of ten P450 genes from the CYP6AE subfamily of *Helicoverpa armigera* involved in xenobiotic metabolism. *Insect Biochemistry and Molecular Biology*. 2018; 93: 79–91. <https://doi.org/10.1016/j.ibmb.2017.12.006> PMID: [29258871](https://pubmed.ncbi.nlm.nih.gov/29258871/)
84. Livak MJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods*. 2001; 25: 402–408. <https://doi.org/10.1006/meth.2001.1262> PMID: [11846609](https://pubmed.ncbi.nlm.nih.gov/11846609/)
85. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*. 1976; 72: 248–254. <https://doi.org/10.1006/abio.1976.9999> PMID: [942051](https://pubmed.ncbi.nlm.nih.gov/942051/)
86. Omura T, Sato R. The carbon monoxide-binding pigment of liver microsomes I Evidence for its hemo-protein nature. *Journal of Biological Chemistry*. 1964; 239: 2370–2378. PMID: [14209971](https://pubmed.ncbi.nlm.nih.gov/14209971/)
87. Xiao H, Ye X, Xu H, Mei Y, Yang Y, Chen X, et al. The genetic adaptations of fall armyworm *Spodoptera frugiperda* facilitated its rapid global dispersal and invasion. *Molecular Ecology Resources*. 2020; 20: 1050–68. <https://doi.org/10.1111/1755-0998.13182> PMID: [32359007](https://pubmed.ncbi.nlm.nih.gov/32359007/)
88. Pearce SL, Clarke DF, East PD, Elfekih S, Gordon KHJ, Jermiin LS, et al. Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC Biology*. 2017; 15: 63. <https://doi.org/10.1186/s12915-017-0402-6> PMID: [28756777](https://pubmed.ncbi.nlm.nih.gov/28756777/)
89. 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 of the United States of America*. 2018; 115: 5034–5039. <https://doi.org/10.1073/pnas.1718831115> PMID: [29610329](https://pubmed.ncbi.nlm.nih.gov/29610329/)