

TOXICOGENOMICS OF *DIABROTICA VIRGIFERA VIRGIFERA* LECONTE IN  
RESPONSE TO ECRY3.1AB-EXPRESSING TRANSGENIC MAIZE

---

A Dissertation  
presented to  
the Faculty of the Graduate School  
at the University of Missouri

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

---

by  
ZIXIAO ZHAO  
Dr. Christine G. Elsik, Dissertation Supervisor

JULY 2019

The undersigned, appointed by the dean of the Graduate School, have examined the  
dissertation entitled

TOXICOGENOMICS OF *DIABROTICA VIRGIFERA VIRGIFERA* LECONTE IN  
RESPONSE TO ECRY3.1AB-EXPRESSING TRANSGENIC MAIZE

Presented by Zixiao Zhao  
a candidate for the degree of  
Doctor of Philosophy

and hereby certify that, in their opinion, it is worthy of acceptance.

---

Advisor, Dr. Christine G. Elsik

---

Co-Advisor, Dr. Bruce E. Hibbard

---

Dr. Kent S. Shelby

---

Dr. J. Chris Pires

## ACKNOWLEDGEMENTS

For decades, generations of scientists have devoted themselves to understand the science behind the western corn rootworm and to limit its impact on agriculture in the U.S. and worldwide. I am very honored to have worked with them and to take it a little step further. Without standing on the shoulders of those scientists, including my Ph.D. committee, I would have not been able to make progress towards my dissertation.

Special thanks must be granted to my advisors Dr. Bruce E. Hibbard and Dr. Christine G. Elsik for their support and guidance during the whole program. The past and current members from Hibbard's lab have accomplished much groundbreaking work on Bt resistance of the western corn rootworm. I thank the many experienced computational biologists in Elsik's lab for software support. Without the support of these groups the dissertation would not have been possible.

I would also like to thank my committee members Dr. Kent S. Shelby and Dr. J. Chris Pires, especially Dr. Shelby who contributed a lot in forming research ideas, implementing confirmation experiments and editing manuscript drafts. Dr. Lisa Meihls initiated this project and offered great help in experimental design and insect dissection. Dr. Tieming Ji reviewed the statistical analysis procedures and provided valuable suggestions. Dr. Adriano Pereira conducted follow-up studies.

Last but not the least thanks to my parents: Yang Zhao and Mei Li. As Karl Marx said, "there is no royal road to science." Their unconditional supports were key for me overcoming the fatigues and fears, continually climbing the steep paths to reach the luminous summits of science in the future.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
ABSTRACT .....	x
<b>Chapter 1. Literature Review</b> .....	1
Introduction.....	2
Diversity of Bt toxins.....	3
Modes of action of Bt $\delta$ -endotoxins.....	5
Sequential binding model .....	6
Signaling pathway model.....	7
ABC transporter model .....	8
The genetics and genomics research of western corn rootworm .....	10
Description of western corn rootworm genome.....	12
Genetic markers of western corn rootworm.....	13
Transcriptome sequencing of western corn rootworm.....	14
Justification and objectives .....	16
Bibliography .....	20
<b>Chapter 2. Differential gene expression in response to eCry3.1Ab ingestion in an unselected and eCry3.1Ab-selected western corn rootworm (<i>Diabrotica virgifera virgifera</i> LeConte)</b> .....	29
Abstract .....	30
Introduction.....	31

Material and methods.....	33
Insects and bioassay .....	33
RNA extraction, library construction and sequencing.....	34
<i>De novo</i> assembly of transcriptome .....	35
Differential expression analysis .....	35
Annotation and pathway analysis .....	36
Results and discussion .....	37
Transcriptome assembly and annotation of WCR transcriptome .....	37
Differential expression analysis of WCR transcriptome .....	38
GO and pathway analysis on eCry3.1Ab feeding WCR midgut .....	39
Expression of potential and novel eCry3.1Ab resistant genes.....	40
Conclusion .....	42
Acknowledgements.....	43
Bibliography .....	55
<b>Chapter 3.</b> The detection of alternative splicing in <i>Diabrotica virgifera virgifera</i>	
LeConte in association with eCry3.1Ab resistance .....	61
Abstract.....	62
Introduction.....	63
Material and methods.....	65
Insects and data source.....	65
Bioassay, Iso-seq library preparation, and sequencing.....	66
Iso-seq data processing and transcript analysis .....	66
Identification of colony-specific alternative spliced genes.....	67

Functional annotation of Iso-seq transcriptome and GO enrichment of colony-specific AS genes.....	68
Characterization of a colony-specific spliced novel peritrophic matrix protein ..	69
Results.....	69
Overview of WCR transcriptome and AS pattern from PacBio Iso-seq .....	69
Functional annotation of WCR Iso-seq transcriptome.....	70
Colony-specific alternative splicing genes of WCR.....	71
Characterization of the colony-specific spliced novel peritrophic matrix protein	72
Discussion.....	74
Acknowledgements.....	77
Bibliography .....	96
<b>Chapter 4. Preliminary analysis of genomic variation in association with eCry3.1Ab</b>	
resistance in western corn rootworm .....	101
Abstract.....	102
Introduction.....	103
Material and methods.....	105
Data source.....	105
RNA-seq reads trimming and alignment to reference genome.....	106
Variant calling, filtering, and principal component analysis .....	107
Variant annotation and intersection analysis .....	108
Results.....	109
Variant calling and PCA analysis .....	109
Variant intersection, annotation and functional analysis .....	109

Discussion.....	111
Bibliography .....	121
<b>VITA</b> .....	126

## LIST OF TABLES

Table	Page
1-1. Comparison of haploid chromosome number and genome size among three model species of Coleopteran insects. ....	18
1-2. The basic statistics of <i>Diabrotica virgifera virgifera</i> genome assembly .....	19
2-1. The design of feeding assay of eCry3.1Ab-resistant and susceptible WCR colony.....	44
2-2. Summary statistics of WCR whole larvae and midgut transcriptome assemblies and their unigene sets.....	45
2-3. Summary of BUSCO analysis of WCR whole larval and midgut transcriptome assemblies and their unigene sets .....	46
2-4. Summary of BLASTX of WCR whole larval and midgut transcriptome assemblies and their unigene sets .....	47
2-3. The top level-2 GO terms of DEG unigene set and midgut transcriptome unigene sets.....	48
3-1. The bioassay treatment of WCR neonate larvae before RNA-extraction and transcriptome PacBio Iso-seq .....	78
3-2. The PCR primers for validate alternative splicing pattern.....	79
3-3. The primer sequences used for qRT-PCR.....	80
3-4. Summary of BUSCO analysis of isoform from Iso-seq .....	81
3-5. The number of isoforms of putative Bt receptor gene cadherin and aminopeptidase N.....	82
3-6. The top enriched GO terms from genes with colony-specific splicing patterns	83
3-7. Two genes with constitutively colony-specific splicing patterns .....	86
4-1. Statistics of detected genomic variants of each sample .....	115
4-2. Statistical parameters of principal component analysis (PCA).....	116
4-3. Pattern of NB-specific variants in potential Bt resistance associated genes....	117

## LIST OF FIGURES

Figure	Page
2-1. Contig length distribution of western corn rootworm larval and midgut transcriptome assemblies .....	50
2-2. The species distribution of BLASTX top hits of WCR larval and midgut transcriptome assemblies .....	51
2-3. Transcriptional response of neonate WCR allowed feed 12 or 24 hrs on eCry3.1Ab or non-Bt isoline maize seedlings compared to resistant neonates feeding on isoline seedlings .....	52
2-4. Differential expression pattern of resistant and susceptible WCR in whole larvae, or dissected midguts, when feeding on Cry3.1Ab transgenic maize seedlings vs. feeding on non-Bt isoline seedlings, for 12 or 24 hrs .....	53
2-5. Expression levels of 11 potential eCry3.1Ab resistant related genes in 8 WCR midgut treatments.....	54
3-1. Functional annotation of isoforms of genes that aligned to WCR genome at level-2 GO terms.....	87
3-2. Venn diagram of genes with colony-specific DEUs under various experimental conditions.....	89
3-3. Venn diagram of enriched GO terms under different condition .....	90
3-4. (A) Diagrammatic representation of gene PBDG00000008671 with its two isoforms; (B) The electrophoresis image of PCR .....	91
3-5. Sequencing alignment of predicted protein sequences from two isoforms and a predicted WCR peritrophic matrix protein C4.....	92
3-6. Signal peptide predictions of PB8671.1 and PB8671.2 by Signal-IP .....	93
3-7. The colony-specific differential exon usage of the novel peritrophic matrix protein gene PBDG00000008671 (alias gene ID: PB8671). .....	94
3-8. The qRT-PCR quantification of novel peritrophic matrix protein PB8671 and isoform PB8671.2 in eCry3.1Ab-susceptible SD and resistance NB colonies..	95
4-1. Plot of principal component analysis of three cohorts of eCry3.1Ab-selected resistant colony (NB) and unselected eCry3.1Ab-susceptible colony (NI).....	118
4-2. Venn diagram of the number of gene loci containing NB (A) and NI (B) specific genetic variations .....	119

4-3. Sequence analysis of a putative WCR ABC transporter subfamily C protein. 120

## ABSTRACT

In modern agricultural systems, insect pests are constantly facing stresses from natural factors and pest management. Strong selection pressure results in evolution of insect resistance to control substances such as chemical insecticides and biological toxins, e.g. *Bacillus thuringiensis* (Bt) toxins. As one of the most severe insect pests for maize in the U.S. Corn Belt, the western corn rootworm (WCR), *Diabrotica virgifera virgifera* LeConte, is in a constant arms race with control strategies. Since 2003, transgenic maize expressing Cry3 toxins have been broadly adopted for rootworm control. WCR has reported resistance to Bt maize in lab-selected colonies and in the field. But the genetics of resistance, and the mode of action of Cry3 toxins to WCR are still not clear. We used eCry3.1Ab-expressing maize and its resistant WCR colony as a model to study Bt resistance. Comprehensive toxicogenomics approaches, including RNA-seq and PacBio Iso-seq, were employed to detect regulation of gene expression regulation, alternative splicing, and genomic variation in eCry3.1Ab-resistance WCR. We also compared the gene expressional response of resistant and susceptible WCR when they were intoxicated with eCry3.1Ab. From these studies we were able to identify genes with functions associated with Bt resistance due to their specific expressional profile, splicing pattern or genomic variation in resistant colonies. We also evaluated differences in responses of resistant and susceptible colonies when being intoxicated by eCry3.1Ab. The preliminary results of genomic variation also reflect the bottleneck of selection process and reveal that a variant allele on a potential Bt receptor increased in the resistant colony. This knowledge will assist our understanding of the unique mode of action of eCry3.1Ab, and develop molecular markers for resistance monitoring in the field.

# **Chapter 1**

## Literature Review

## Introduction

*Bacillus thuringiensis* Berliner (Bt) is a gram-positive, spore-forming soil entomopathogenic bacterium that belongs to the order Bacillales. It was independently isolated from infected silkworm (*Bombyx mori*) and Mediterranean flour moth (*Ephestia kuehniella*) in the early 20<sup>th</sup> century (Beegle and Yamamoto 1992; Milner 1994). Later, the toxicity of Bt was confirmed from the parasporal protein crystal of an inclusion body alongside the spore, which is unique in Bt (Angus 1956; Hannay 1953; Hannay and Fitz-James 1955; Heimpel and Angus 1959). The protein toxins identified in the parasporal crystal were later named as  $\delta$ -endotoxins (Heimpel 1967). Two other types of Bt entomotoxins, named vegetative insecticidal proteins (Vip) (Estruch et al. 1996) and secreted insecticidal protein (Sip) (Donovan et al. 2006), were discovered in late 20<sup>th</sup> and early 21<sup>st</sup> centuries.

Since its discovery, Bt has been adopted as a biocontrol agent. Sporeine was the first Bt product and was first commercialized in France in 1938 (Beegle and Yamamoto 1992; Milner 1994). In the 1950's, Bt biopesticides were widely used in Europe and the US. Four *B. thuringiensis* subspecies were registered with the Environmental Protection Agency (EPA) and commercialized for their effectiveness against insect pests: subsp. *kurstaki* comprises the largest portion of the Bt biopesticide market for controlling multiple Lepidopteran species in vegetable production; subsp. *aizawai* specifically targets *Spodoptera* spp. and wax moth (*Galleria mellonella*) in honey comb; subsp. *israelensis* targets Dipteran larvae and is used for controlling mosquitos and blackflies; subsp. *tenebrionis* (a.k.a. subsp. *san diego*) targets Coleopteran species and is used for controlling elm leaf beetle (*Xanthogaleruca luteola* (Müller)) and Colorado potato beetle

(*Leptinotarsa decemlineata* Say) (Beegle and Yamamoto 1992). Formulations of those biopesticides are based on live Bt spores, which have the disadvantage of short period of effectiveness because crystals quickly break down under the sunlight causing loss of toxicity (Pusztai et al. 1991). The effectiveness of Bt pesticides is also affected by environmental conditions, i.e. proteases in leaf exudates, rainfall and wind (Sanahuja et al. 2011).

Biotechnology has allowed plants to gain toxicity against insect pests by transforming genes encoding Bt toxins into crop genomes. Since 1996, genes expressing Bt or engineered Bt toxins have been introduced into major row crops including maize, soybean, cotton, and other crops. The diversity of Bt toxins has facilitated the search for new Bt molecules for battling against previously uncontrolled insect pests. Due to their effectiveness, worldwide planting of Bt-crops has been extensive and also imposes selection pressures, which has led to the problem of Bt resistance in insect pests (Tabashnik and Carrière 2017). This review focuses on toxin structures, recent discoveries of modes of action, and the genetics of western corn rootworm (*Diabrotica virgifera virgifera* LeConte), of which the Bt resistance will be studied in following chapters.

### **Diversity of Bt toxins**

The first  $\delta$ -endotoxin genes were cloned by Schnepf & Whiteley (1981) from a large plasmid of *B. t.* subsp. *kurstaki* strain HD-1. Höfte & Whiteley (1989) first categorized the growing numbers of  $\delta$ -endotoxin genes into four *cry* classes and one *cyt* class based on sequence homology and target spectrum of protein: *cryI* targets Lepidoptera, *cryII* targets Lepidoptera and Diptera, *cryIII* targets Coleoptera, and *cryIV*

targets Diptera. The *cytA* gene, which encodes a 27-kDa protein, was discovered from *B. t. subsp. israelensis*. It does not have sequence homology with *cry* classes but shows toxicity to Diptera and some vertebrate and invertebrate cell.

The first nomenclature quickly became overwhelmed when new toxins were continuously discovered and could not fit into any class. A new hierarchical nomenclature based on molecular evolution history was accepted (Crickmore et al. 1998). The *cry/cyt* gene superfamily names were kept. However, instead of being assigned into classes, the genes were given a primary rank (numeric number), a secondary rank (upper case letter), a tertiary rank (lower case letter) and a quaternary rank (a numeric number) if necessary, solely based on sequence homology. To date, more than 800  $\delta$ -endotoxins have been documented and clustered into 75 Cry families (Cry1 to Cry75) and 3 Cyt families (Cyt1 to Cyt3) (Crickmore et al. 2018; Palma et al. 2014; Xiao and Wu 2019). Some Cry toxins do not have any invertebrate target and they are referred as parasporins (Okumura et al. 2010; Palma et al. 2014).

Cry toxins can be further categorized into three types based on protein structure. The majorities have molecular weight of 65 or 130 kDa and feature three conservative domains with similar three-dimensional structures. They are categorized as three-domain Cry (3d-Cry) toxins (Pardo-López et al. 2013; Sanahuja et al. 2011; Schnepf et al. 1998; Xu et al. 2014). The second type is ETX/MTX-like. Proteins of this type are featured with ETX/MTX domain that similar to epsilon toxins (ETX) from bacterium *Clostridium perfringens* (formally *Bacillus welchii*) and mosquitocidal toxins (MTX) from bacterium *Lysinibacillus sphaericus* (formerly *Bacillus sphaericus*) (de Maagd et al. 2003; Xu et al. 2014). The third type is binary-like. Toxins from this type usually consist of two separate

molecules. Binary toxins were first found in *L. sphaericus* (de Maagd et al. 2003). In Bt, binary pairs were found between Cry34 and Cry35, Cry48 and Cry49, and Cry37Aa1 with Cry23Aa (Crickmore et al. 2018; de Maagd et al. 2003). Both sequence divergence and domain swapping contribute to structure diversity, and expand the spectrum of target specificity to arthropods (de Maagd et al. 2001).

Vip and Sip are two types of Bt entomotoxins that do not form parasporal crystals. The *vip* genes are expressed in the vegetative growth stage (Estruch et al. 1996). A similar nomenclature is adopted for Vip and Sip. To date, four Vip families were discovered, ranging from Vip1 to Vip4 (Crickmore et al. 2018). Vip1 and Vip2 form binary toxins that resemble to Cry binary toxins. The target spectrum includes some Hemipteran, Coleopteran, and Dipteran species (Chakroun et al. 2016). Sip is a novel toxin discovered in the culture supernatant of wild-type *B. t.* strains. The only Sip protein characterized as targeting Coleopteran species was Sip1A, which has toxicity on Colorado potato beetle, southern corn rootworm (*Diabrotica undecimpunctata howardi*), western corn rootworm, and cabbage beetle (*Colaphellus bowringi*) (Donovan et al. 2006; Sha et al. 2018).

### **Mode of action of *Bt* $\delta$ -endotoxins**

Research on Bt-insect interactions has been mainly focused on 3d-Cry toxins due to their broad applications in biopesticides and biotechnology. The toxin enters the insect digestive tract by feeding. Parasporal crystals are solubilized in midgut lumen under alkalic pH and reducing atmosphere, releasing the protoxin. In transgenic plants, the protoxins are soluble and can be directly taken up by insects. After uptake, parts of the N- and C-terminal regions are removed by trypsin or other gut proteases, yielding a 45-70

kDa active toxin (Bravo et al. 2007; Rukmini et al. 2000). The active toxin then passes through the midgut peritrophic membrane, and once in contact with midgut epithelium cells, causes cell death, which interrupts gut functions and causes. The details of this process are not yet fully elucidated. The most intensive research in this area has been done on the Lepidopteran insect tobacco hornworm (*Manduca sexta*) with Cry1A toxins. Two “classic” models were proposed based on genetics and biochemical studies: sequential binding model and signaling pathway model (Melo et al. 2016). Vachon et al. (2012) reviewed those models and evaluated the evidence and exceptions of each model.

Recently, much attention has been given to ATP-binding cassette transporter (ABC transporter), especially families C, A and B. The roles of ABC transporters in Bt mode of action are not fully elucidated. There are hypotheses supported with evidence that ABC transporters might be involved in the process of sequential binding model (Heckel 2012; Ocelotl et al. 2017), while in other cases, they can have independent impacts on Bt toxicity (Tanaka et al. 2017). Here we suggest a third model: the ABC transporter model. The three models are not mutually exclusive and it is possible that Cry toxin incorporates more than one model to achieve toxicity towards epithelium cell toxicity and there might be more models yet to be discovered.

### **Sequential binding model**

The sequential binding model has been extensively studied and reviewed (Bravo et al. 2007; Pardo-López et al. 2013). In this model, cell toxicity is accomplished by formation of non-selective channels by *Bt* toxin, causing increased permeability, loss of membrane function, and death of the cells. The pore formation process is mediated by many membrane receptors. The monomeric toxins first bind to glycosyl

phosphatidylinositol (GPL) anchored alkaline phosphatase (ALP) or aminopeptidase N (APN). Both APN and ALP have low affinity to the monomeric toxins, but their abundances are high in the cell membrane. Their first role is to concentrate the monomeric toxins, allowing the toxins to bind to the low abundance, high-affinity cadherin-like protein (CAD). CAD further proteolyzes the toxin at its N-terminal end, removing helix  $\alpha$ -1, and allows the remaining peptide to form a pre-pore oligomer. The affinity of pre-pore oligomer to ALP and APN increases 200-fold compared to the monomeric form (Bravo et al. 2004). The pre-pore oligomer binds to ALP and APN, triggers its conformational change and inserts into the cell membrane lipid raft, resulting in pore formation and eventually, the death of the midgut epithelial cells (Bravo et al. 2007; Pardo-López et al. 2013).

### **Signaling pathway model**

The signaling pathway model focuses on programmed cell death induced by Cry toxins. The first case was reported in common house mosquito (*Culex pipiens*) larvae where Cry toxin from *B. t.* subsp. *israelensis* strain HD522 can induce the apoptosis of their midgut cells in both *in vivo* and *in vitro* studies (Smouse and Nishiura 1997). The Cry toxins of this strain are Cry4Aa2 or Cry4Ba4, according to Bt toxin nomenclature database (Crickmore et al. 2018). *In vivo* studies on *Bombyx mori* showed Cry1Aa can induce apoptosis of midgut cells, but only under nonlethal and sublethal conditions (Tanaka et al. 2012).

A series of papers were published based on studies of Cry1Ab with its cadherin receptor Bt-R1 identified in *M. sexta* (Zhang et al. 2005; Zhang et al. 2006; Zhang et al. 2008). Programmed cell death was observed in the cabbage looper (*Trichoplusia ni*) H5

cell line expressing heterologous Bt-R1 when challenged by Cry1Ab. These studies showed that cadherin binding with activated toxin stimulates G proteins, which activates adenylyl cyclase (AC), causing an increased level of cAMP, the intracellular secondary messenger. The cAMP binds to protein kinase A (PKA) to initiate a series of signaling cascades, resulting in oncosis of the cells. The AC/PKA pathway also triggers the exocytotic transport of cadherin receptor, which amplifies the signal (Zhang et al. 2008). This process is  $Mg^{2+}$ -dependent and does not require oligomerization of toxin.

The signaling pathway model provides a different scope of intracellular signaling of Cry toxicity. However, the activated signaling pathways and their effects seem to be unique for each Bt toxin. The studies on eastern spruce budworm (*Choristoneura fumiferana*) CF-1 cell show Cry1Ab and Cry1Ac do not activate AC/PKA pathway. Instead, they activate MAPK pathway and trigger cell apoptosis. Pore formation is also necessary for this process (Portugal et al. 2017). Additionally, Cry6Aa, a non-3D toxin, confers nematocidal activity to *Caenorhabditis elegans* by triggering  $Ca^{2+}$  dependent necrosis pathway (Zhang et al. 2016). Although the signaling pathway models were supported by cultured cell assays, no resistant insect phenotype was found yet to support the model.

### **ABC transporter model**

It was not until the first decade of the 21<sup>st</sup> century that ABC transporters were found linked with Bt resistance by forward genetics studies. Gahan et al. (2010) first reported that a mutation of an ABCC2 transporter causes resistance to Cry1Ac in tobacco budworm (*Heliothis virescens*). Later, cases of Bt resistance associated with ABC transporters were reported in many Lepidopteran species including diamondback moth

(*Plutella xylostella*) (Guo et al. 2015), silkworm (Atsumi et al. 2012), corn earworm (*Helicoverpa armigera*) (Xiao et al. 2014), fall armyworm (*Spodoptera frugiperda*) (Banerjee et al. 2017), pink bollworm (*Pectinophora gossypiella*) (Mathew et al. 2018), and Asian corn borer (*Ostrinia furnacalis*) (Coates and Siegfried 2015; Zhang et al. 2017).

In order to fit the role of ABC transport to the sequential binding model, Heckel (2012) proposed that ABC transporters might (1) connect with pre-pore toxin oligomers and assist the insertion and stabilization of the pore structure and (2) assist oligomerization. This hypothesis is supported by the studies of Cry1AMod toxins, which were derived from Cry1A by removing their N-terminal helix  $\alpha$ -1. The Cry1AMod mimics the products of cadherin process and forms pre-pore structures in the absence of cadherin (Soberón et al. 2007). The *P. xylostella* Cry1A-resistant strain NO-QAGE has a mutation on the ABCC2 transporter. The insertion of Cry1AMod pre-pore structure in NO-QAGE is less efficient than in the wild type strain (Ocelotl et al. 2017). The result suggested that ABCC2 has effects at the downstream of cadherin processes, and the insertion would be irreversible without a functional ABCC2.

Many recent studies of *Bombyx mori* ABCC2 (BmABCC2) in cultured cell systems have supported that an ABC transporter could be a primary receptor of Cry toxins. ABC transporter-mediated pore formation does not require cadherin, ALP, or APN. BmABCC2 can directly bind to Cry1A toxins with its extracellular loop 4 (ECL4) with high affinity (Tanaka et al. 2017). Expression of BmABCC2 allows Cry1Aa to form stable pores, which dramatically increases its susceptibility. Expressing *Bombyx mori* cadherin BtR also increases susceptibility, but to a lower degree (Tanaka et al. 2016;

Tanaka et al. 2013). The results suggest that endogenous cadherin of Sf9 might not be sufficient to activate Cry1A. Moreover, increased susceptibility to Cry1Aa was also found in human cell line HEK293T and *Xenopus* oocyte when BmABCC2 was introduced and expressed (Endo et al. 2017). These results further support that ABCC might mediate an independent mode of action (Endo et al. 2017). Synergistic effects were observed between ABCC2 and cadherin (Tanaka et al. 2016; Tanaka et al. 2013), suggesting that Cry toxins may use more than one model and there may be cross talk between models.

Cell culture systems have been used to study other ABC transporters from several species (Sato et al. 2019). Lepidopteran ABCC transporters, including BmABCC2, BmABCC3 and SeABCC2 and SeABCC3 from beet armyworm (*Spodoptera exigua*) are confirmed receptors of Cry1Aa and Cry1Ab but not Cry1Ca and Cry1Da. Some ABC transporters can also serve as receptors for different Cry families. For example, BmABCC2 and SeABCC3 also functioned as receptors of Cry8Ca, a toxin targeting Coleoptera. Furthermore, CtABCB1 from a leaf beetle (*Chrysomela tremula*) and BmABCA2 show receptor function with Cry3Aa and Cry2Ab respectively (Endo et al. 2017; Sato et al. 2019). These results suggest that ABC transporters might be one of the determinants of Cry target specificity.

### **Genetic and genomic research of western corn rootworm**

Rootworm is a group of *Diabrotica* spp. in the order of Chrysomelidae (Order: Coleoptera). There are four rootworms: northern corn rootworm (NCR, *D. barberi*), southern corn rootworm (SCR, *D. undecimpunctata howardi*), western corn rootworm (WCR, *D. virgifera virgifera*), and Mexican corn rootworm (MCR, *D. virgifera zea*). In

the U.S., only NCR and WCR cause significant annual damage to maize production. It is estimated that before transgenic maize was applied, the combined cost of control and yield loss of both rootworms was more than \$1 billion annually in the U.S. (Gray et al. 2009). WCR is notorious for its adaptation to control strategies. It has evolved resistance to many control strategies including broadcast soil insecticides (Ball and Weekman 1962), crop rotation (Levine et al. 2002), transgenic maize expressing Cry toxins (Gassmann et al. 2011; Ludwick and Hibbard 2016; Meihls et al. 2008; Zukoff et al. 2016) and even RNAi transgenic maize expressing double stranded RNA for rootworm control (Khajuria et al. 2018).

The WCR is also an invasive species. The initial distribution was limited primarily to western Kansas and eastern Colorado (Chiang 1973) although remnant populations can also be found in Mexico and likely other areas of southwest U.S. (Segura-León 2004). After the adoption of continuous monoculture maize production, the WCR had a rapid eastward expansions across the Great Plains and the current Corn Belt (Gray et al. 2009). In 1992, the first WCR in the Europe was found near the Belgrade Airport of Yugoslavia (now Serbia). Since then, the WCR has expanded its range in Europe to include more than 20 countries with multiple introduction events (Ciosi et al. 2008; Miller et al. 2005). It is estimated that the potential range of WCR overlaps with 64% of global maize production area (Kriticos et al. 2012), which poses strong risks on global food security and raises international concern.

Research on genetics and genomics of WCR allows better understanding of the genetic basis of adaptation, and leads to improved strategies for monitoring resistance to *Bt* and insecticides, predicting worldwide expansion and developing new control

strategies (Miller et al. 2009; Sappington et al. 2006). To enhance the collaboration on rootworm genetics research and genome sequencing project, the *Diabrotica* Genetics Consortium was formed in 2003 (Miller et al. 2010; Sappington 2011). It was not until 15 years later that the first version of *D. v. virgifera* draft genome assembly was available to the public (<https://www.ncbi.nlm.nih.gov/genome/?term=diabrotica>). During that period, much research was done using advanced techniques and much preliminary data required for genome sequencing was generated.

### **Description of WCR genome**

WCR has a chromosome number of  $2n=18+XO$  (Ennis 1972). The primary size estimation of its haploid genome is 2.5 Gbp (Miller et al. 2010; Sappington et al. 2006). Flow cytometry assay provides a better estimation of 2.58 Gbp (Coates et al. 2012), which is comparable with mice (2.8 Gbp), humans (3.2 Gbp) and maize (2.5 Gbp). The genome size of WCR is significantly larger than two other Coleoptera species: red flour beetle (*Tribolium castaneum*) and Colorado potato beetle, albeit WCR has the same number of chromosomes as *T. castaneum* (Table 1-1. data source: Hsiao and Hsiao 1983; Richards et al. 2008; Schoville et al. 2018). The genome contains high proportions of repetitive DNA and transposable elements (TE). Many of these TEs are retroelements and located in intergenic regions. The movement of retroelements might contribute to genomic variation, thus plasticity to environment (Coates et al. 2012). However, the high percentage of TEs and repetitive DNA imposes significant hurdles in genome assembly.

The current version of the western corn rootworm genome scaffold assembly (version 2, [https://www.ncbi.nlm.nih.gov/assembly/GCF\\_003013835.1/](https://www.ncbi.nlm.nih.gov/assembly/GCF_003013835.1/)) is sequenced from a non-diapause inbred strain Ped12-6-A-3. About 2.41 Gbp of nucleotides are

assembled into the current genome assembly. Compared to the initial version, the number of scaffolds is reduced by 50%, indicating increased continuity (Table 1-2). However a large number of transcriptome PacBio Iso-seq reads failed to align to the current genome assembly, suggesting that a significant proportion of coding region is not covered (see Chapter 3).

### **Genetic markers of western corn rootworm**

Genetic markers are powerful tools to analyze population structure and to map biologically and economically important traits to the genome. Preliminary genetic markers of WCR and related research have been reviewed by Miller et al. (2009). Many markers were discovered in sequencing projects (i.e. restriction enzyme fragment sequencing, expressed sequence tag (EST) sequencing, and transcriptome sequencing via second-generation high-throughput RNA-sequencing (RNA-seq) projects). The development of markers as well as the sequencing project together paved the way towards genetic knowledge of WCR.

The first applicable markers were microsatellites (Kim and Sappington 2005b). Microsatellites, or simple sequence repeats (SSR), are tandem repeats of DNA sequences that are usually polymorphic. Kim et al. (2008) recommended a core set of 22 SSR markers for their high polymorphism, unambiguous readability and repeatability, selective neutrality, lack of non-alleles, and no linkage between loci. In western corn rootworm, SSR markers have been used to study population structures within the US (Kim and Sappington 2005a), adaptation to maize-soybean rotation (Miller et al. 2006), and to reveal the introduction events in Europe (Miller et al. 2005).

Single nucleotide polymorphisms (SNPs) are each caused by mutation of a single base within the genome. SNPs are highly abundant in the rootworm genome, which makes them particularly suitable for gene mapping, genotyping, and resistance monitoring purposes. Nonsynonymous SNPs, which are SNPs in coding regions that change the amino acid sequence of protein, may result in altered protein function with phenotypical consequences (Wang et al. 2013). The first sets of western corn rootworm SNPs markers were based on EST sequencing data (Coates et al. 2009). About 4000 SNPs were detected and used for mapping traits (Coates et al. 2016). Advances in sequencing technology have greatly accelerated the discovery of markers. Flagel et al. (2014) applied a comprehensive approach in combination of second-generation 454 sequencing as reference and Illumina DNA-sequencing for 26 populations. Over 500,000 SNPs were generated in the gene region. With high-throughput genotyping methods it is now feasible to map quantitative traits such as Cry3Bb1 resistance to linkage groups and find the candidate genes related with resistance (Flagel et al. 2015). However, the current SNP markers are located in transcribed regions. SNPs of intron and intergenic regions that might reveal impact of regulatory mechanisms in western corn rootworm have yet to be studied.

### **Transcriptome sequencing of western corn rootworm**

Transcriptome sequencing provides insights into biological mechanisms. Before second-generation sequencing was invented, expressed transcripts of western corn rootworm were sequenced using first-generation Sanger sequencing methods. The first EST dataset contained only 691 sequences, and laid the foundation for rootworm genomic and transcriptomic research (Siegfried et al. 2005). Since then, the NCBI dbEST

database has collected 17,796 sequences from WCR. The EST sequences also served as reference in a variety projects, including microarray studies of rotation resistance and SNP analysis of organophosphate resistance traits (Coates et al. 2016; Knolhoff et al. 2010).

Second-generation sequencing technologies provide cost effective methods to generate high-throughput transcriptome sequencing data. For the relatively higher read-length, 454 pyrosequencing continues to be used for EST sequencing while RNA-seq using Illumina technology generates high coverage short reads, allowing the analysis of sequence information and quantification of expression level at the same time (Metzker 2010). For WCR, the transcriptome must be *de novo* sequenced and assembled from either 454 (Flagel et al. 2014), RNA-seq (Miller and Zhao 2015; Zhao et al. 2019), or a combination of the two (Chu et al. 2015; Eyun et al. 2014).

In parallel to genomic studies using markers, many RNA-seq studies on WCR focus on gene expression regulation associated with adaptations to plant defense chemicals and control strategies. An RNA-seq study demonstrated that the maize hydroxamic acid DIMBOA is able to induce the up-regulation of five transcripts, including putative cytochrome P450 and cathepsin proteases, which implies that inducible defense is involved in adaption to plant defense chemicals (Miller and Zhao 2015). Another RNA-seq study on rotation resistant and susceptible western corn rootworm adults revealed a divergence of transcriptome profile between populations (Chu et al. 2015).

With the widespread use of Bt-maize expressing rootworm-target Cry toxins, the western corn rootworm has developed field resistance to Cry3Bb1 (Gassmann et al. 2011;

Schrader et al. 2017; Wangila et al. 2015). Cross-resistance among three toxins (Cry3Bb1, mCry3A and eCry3.1Ab) from Cry3 families has been reported (Zukoff et al. 2016). Incomplete field resistant populations to Cry34/35Ab1 were reported, with no cross-resistance with Cry3 toxins (Gassmann et al. 2016; Ludwick et al. 2017). However, the detailed mode of action of Cry34/35Ab1 and Cry3 family toxins to WCR has not been elucidated. RNA-seq has been extensively employed to explore the modes of action and resistant mechanisms of Cry3Bb1 (Rault et al. 2018), eCry3.1Ab (Zhao et al. 2019), and Cry34/35Ab1 (Wang et al. 2017). Both Cry3Bb1 and eCry3.1Ab induce dramatic changes in expression profiles of susceptible larvae, but the not resistant ones. Wang et al. (2017) did not have access to Cry34/35Ab-resistant WCR colonies. Although WCR cadherin shows binding capability in *in vitro* assay (Park et al. 2009), it was not differentially expressed between resistant and susceptible WCR. Both Cry3Bb1 and eCry3.1Ab induce up-regulation of ABC transporters in WCR especially in susceptible larvae. The results indicate ABC transporters might have important roles in Cry3 mode of action and mechanism of resistance in western corn rootworm.

### **Justifications and objectives**

The eCry3.1Ab protein is a hybrid toxin consisting of an mCry3A backbone and a C-terminal domain from Cry1Ab (Walters et al. 2010). It is the fourth commercialized *Bt* toxin targeting corn rootworm. Resistance to eCry3.1Ab was developed through a laboratory selection experiment (Frank et al. 2013) and resistance to eCry3.1Ab is also present as cross-resistance in some field strains (Zukoff et al. 2016). The genetic basis of eCry3.1Ab mode of action and resistance mechanism has not previously been studied. The overall objectives of this study is to (1) identify the regulation of gene expression in

western corn rootworm in response to eCry3.1Ab intoxication, and the expressional regulation between susceptible and resistant population, and (2) identify the effects of a post-transcriptional process, alternative splicing, related to eCry3.1Ab resistance. We also carried out (3) a preliminary SNP discovery associated with eCry3.1Ab resistance. The research provided pipelines and formed a framework of comprehensive study of insect resistance using various sequencing tools. The results will contribute to further understanding of mode of action of Cry3 toxins, developing markers for resistant monitoring, and leading to invention of novel rootworm-targeting toxins.

**Table 1-1.** Comparison of haploid chromosome number and genome size among three Coleopteran species: western corn rootworm (*Diabrotica virgifera virgifera*), red flour beetle (*Tribolium castaneum*) and Colorado potato beetle (*Leptinotarsa decemlineata*)

<b>Species</b>	<b>Number of haploid chromosome</b>	<b>Genome size (Mb)</b>
<i>Diabrotica virgifera virgifera</i>	10	2,580
<i>Tribolium castaneum</i>	10	151
<i>Leptinotarsa decemlineata</i>	18	1,170

**Table 1-2.** The basic statistics of *Diabrotica virgifera virgifera* genome assembly.

Version 2 is the currently assembly published on October 2018. Version 1 was published on March 2018.

	<b>Version 2</b>	<b>Version1</b>
Total sequence length	2,418,073,815	2,409,977,489
Total ungapped length	1,850,692,199	1,850,759,093
Number of scaffolds	87,712	164,630
Scaffold N50	489,108	95,089
Scaffold L50	1,128	6,022
Number of contigs	585,680	579,519
Contig N50	6,238	6,289
Contig L50	84,252	83,799

## Bibliography

- Angus, T. A. (1956) Association of toxicity with protein-crystalline inclusions of *Bacillus sotto* Ishiwata. *Canadian Journal of Microbiology* **2**: 122-131.
- Atsumi, S., Miyamoto, K., Yamamoto, K., Narukawa, J., Kawai, S., Sezutsu, H., *et al.* (2012) Single amino acid mutation in an ATP-binding cassette transporter gene causes resistance to Bt toxin Cry1Ab in the silkworm, *Bombyx mori*. *Proceedings of the National Academy of Sciences* **109**: E1591-E1598.
- Ball, H. J. and Weekman, G. T. (1962) Insecticide resistance in the adult western corn rootworm in Nebraska. *Journal of Economic Entomology* **55**: 439-441.
- Banerjee, R., Hasler, J., Meagher, R., Nagoshi, R., Hietala, L., Huang, F., *et al.* (2017) Mechanism and DNA-based detection of field-evolved resistance to transgenic Bt corn in fall armyworm (*Spodoptera frugiperda*). *Scientific Reports* **7**: 10877.
- Beegle, C. C. and Yamamoto, T. (1992) History of *Bacillus thuringiensis* Berliner research and development. *The Canadian Entomologist* **124**: 587-616.
- Bravo, A., Gill, S. S. and Soberón, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49**: 423-435.
- Bravo, A., Gómez, I., Conde, J., Muñoz-Garay, C., Sánchez, J., Miranda, R., *et al.* (2004) Oligomerization triggers binding of a *Bacillus thuringiensis* Cry1Ab pore-forming toxin to aminopeptidase N receptor leading to insertion into membrane microdomains. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1667**: 38-46.
- Chakroun, M., Banyuls, N., Bel, Y., Escriche, B. and Ferré, J. (2016) Bacterial vegetative insecticidal proteins (Vip) from entomopathogenic bacteria. *Microbiology and Molecular Biology Reviews* **80**: 329-350.
- Chiang, H. C. (1973) Bionomics of the northern and western corn rootworms. *Annual Review of Entomology* **18**: 47-72.
- Chu, C.-C., Zavala, J. A., Spencer, J. L., Curzi, M. J., Fields, C. J., Drnevich, J., *et al.* (2015) Patterns of differential gene expression in adult rotation-resistant and wild-type western corn rootworm digestive tracts. *Evolutionary Applications* **8**: 692-704.
- Ciosi, M., Miller, N. J., Kim, K. S., Giordano, R., Estoup, A. and Guillemaud, T. (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology* **17**: 3614-3627.
- Coates, B. S., Alves, A. P., Wang, H., Walden, K. K., French, B. W., Miller, N. J., *et al.* (2012) Distribution of genes and repetitive elements in the *Diabrotica virgifera virgifera* genome estimated using BAC sequencing. *Journal of Biomedicine and Biotechnology* **2012**.

- Coates, B. S., Alves, A. P., Wang, H., Zhou, X., Nowatzki, T. M., Chen, H., *et al.* (2016) Quantitative trait locus mapping and functional genomics of an organophosphate resistance trait in the western corn rootworm, *Diabrotica virgifera virgifera*. *Insect Molecular Biology* **25**: 1-15.
- Coates, B. S. and Siegfried, B. D. (2015) Linkage of an ABCC transporter to a single QTL that controls *Ostrinia nubilalis* larval resistance to the *Bacillus thuringiensis* Cry1Fa toxin. *Insect Biochemistry and Molecular Biology* **63**: 86-96.
- Coates, B. S., Sumerford, D. V., Miller, N. J., Kim, K. S., Sappington, T. W., Siegfried, B. D., *et al.* (2009) Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity* **100**: 556-564.
- Crickmore, N., Baum, J. A., Bravo, A., Lereclus, D., Narva, K., Sampson, K., *et al.* (2018) *Bacillus thuringiensis* toxin nomenclature. <http://www.btnomenclature.info/>.
- Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, H. E., Van Rie, J., Lereclus, D., *et al.* (1998) Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews* **62**: 807-813.
- de Maagd, R. A., Bravo, A., Berry, C., Crickmore, N. and Schnepf, H. E. (2003) Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annual Review of Genetics* **37**: 409-433.
- de Maagd, R. A., Bravo, A. and Crickmore, N. (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends in Genetics* **17**: 193-199.
- Donovan, W. P., Engleman, J. T., Donovan, J. C., Baum, J. A., Bunkers, G. J., Chi, D. J., *et al.* (2006) Discovery and characterization of Sip1A: a novel secreted protein from *Bacillus thuringiensis* with activity against coleopteran larvae. *Applied Microbiology and Biotechnology* **72**: 713-719.
- Endo, H., Tanaka, S., Imamura, K., Adegawa, S., Kikuta, S. and Sato, R. (2017) Cry toxin specificities of insect ABCC transporters closely related to lepidopteran ABCC2 transporters. *Peptides* **98**: 86-92.
- Ennis, T. (1972) Meiosis in *Diabrotica* (Coleoptera: Chrysomelidae): chiasma frequency and variation. *Canadian Journal of Genetics and Cytology* **14**: 113-128.
- Estruch, J. J., Warren, G. W., Mullins, M. A., Nye, G. J., Craig, J. A. and Koziel, M. G. (1996) Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proceedings of the National Academy of Sciences* **93**: 5389-5394.
- Eyun, S.-i., Wang, H., Pauchet, Y., French-Constant, R. H., Benson, A. K., Valencia-Jiménez, A., *et al.* (2014) Molecular evolution of glycoside hydrolase genes in the western corn rootworm (*Diabrotica virgifera virgifera*). *PLoS ONE* **9**: e94052.

- Flagel, L. E., Bansal, R., Kerstetter, R. A., Chen, M., Carroll, M., Flannagan, R., *et al.* (2014) Western corn rootworm (*Diabrotica virgifera virgifera*) transcriptome assembly and genomic analysis of population structure. *BMC Genomics* **15**: 1-13.
- Flagel, L. E., Swarup, S., Chen, M., Bauer, C., Wanjugi, H., Carroll, M., *et al.* (2015) Genetic markers for western corn rootworm resistance to Bt toxin. *G3: Genes|Genomes|Genetics* **5**: 399-405.
- Frank, D. L., Zukoff, A., Barry, J., Higdon, M. L. and Hibbard, B. E. (2013) Development of resistance to eCry3.1Ab-expressing transgenic maize in a laboratory-selected population of western corn rootworm (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **106**: 2506-2513.
- Gahan, L. J., Pauchet, Y., Vogel, H. and Heckel, D. G. (2010) An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLOS Genetics* **6**: e1001248.
- Gassmann, A. J., Petzold-Maxwell, J. L., Keweshan, R. S. and Dunbar, M. W. (2011) Field-evolved resistance to Bt maize by western corn rootworm. *PLoS One* **6**: e22629.
- Gassmann, A. J., Shrestha, R. B., Jakka, S. R. K., Dunbar, M. W., Clifton, E. H., Paolino, A. R., *et al.* (2016) Evidence of resistance to Cry34/35Ab1 corn by western corn rootworm (Coleoptera: Chrysomelidae): root injury in the field and larval survival in plant-based bioassays. *Journal of Economic Entomology* **109**: 1872-1880.
- Gray, M. E., Sappington, T. W., Miller, N. J., Moeser, J. and Bohn, M. O. (2009) Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annual Review of Entomology* **54**: 303-321.
- Guo, Z., Kang, S., Zhu, X., Xia, J., Wu, Q., Wang, S., *et al.* (2015) Down-regulation of a novel ABC transporter gene (*Pxwhite*) is associated with Cry1Ac resistance in the diamondback moth, *Plutella xylostella* (L.). *Insect Biochemistry and Molecular Biology* **59**: 30-40.
- Hannay, C. L. (1953) Crystalline inclusions in aerobic sporeforming bacteria. *Nature* **172**: 1004.
- Hannay, C. L. and Fitz-James, P. (1955) The protein crystals of *Bacillus thuringiensis* Berliner. *Canadian Journal of Microbiology* **1**: 694-710.
- Heckel, D. G. (2012) Learning the ABCs of Bt: ABC transporters and insect resistance to *Bacillus thuringiensis* provide clues to a crucial step in toxin mode of action. *Pesticide Biochemistry and Physiology* **104**: 103-110.
- Heimpel, A. M. (1967) A critical review of *Bacillus thuringiensis* var. *thuringiensis* Berliner and other crystalliferous bacteria. *Annual Review of Entomology* **12**: 287-322.

- Heimpel, A. M. and Angus, T. A. (1959) The site of action of crystalliferous bacteria in lepidoptera larvae. *Journal of Insect Pathology* **1**: 152-170.
- Höfte, H. and Whiteley, H. R. (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiological Reviews* **53**: 242-255.
- Hsiao, T. H. and Hsiao, C. (1983) Chromosomal analysis of *Leptinotarsa* and *Labidomera* species (Coleoptera: Chrysomelidae). *Genetica* **60**: 139-150.
- Khajuria, C., Ivashuta, S., Wiggins, E., Flagel, L., Moar, W., Pleau, M., *et al.* (2018) Development and characterization of the first dsRNA-resistant insect population from western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *PLOS ONE* **13**: e0197059.
- Kim, K. S. and Sappington, T. W. (2005a) Genetic structuring of western corn rootworm (Coleoptera: Chrysomelidae) populations in the United States based on microsatellite loci analysis. *Environmental Entomology* **34**: 494-503.
- Kim, K. S. and Sappington, T. W. (2005b) Polymorphic microsatellite loci from the western corn rootworm (Insecta: Coleoptera: Chrysomelidae) and cross-amplification with other *Diabrotica* spp. *Molecular Ecology Notes* **5**: 115-117.
- Kim, K. S., Stolz, U., Miller, N. J., Waits, E. R., Guillemaud, T., Sumerford, D. V., *et al.* (2008) A core set of microsatellite markers for western corn rootworm (Coleoptera: Chrysomelidae) population genetics studies. *Environmental Entomology* **37**: 293-300.
- Knolhoff, L. M., Walden, K. K. O., Ratcliffe, S. T., Onstad, D. W. and Robertson, H. M. (2010) Microarray analysis yields candidate markers for rotation resistance in the western corn rootworm beetle, *Diabrotica virgifera virgifera*. *Evolutionary Applications* **3**: 17-27.
- Kriticos, D. J., Reynaud, P., Baker, R. H. A. and Eyre, D. P. (2012) Estimating the global area of potential establishment for the western corn rootworm (*Diabrotica virgifera virgifera*) under rain-fed and irrigated agriculture. *EPPO Bulletin* **42**: 56-64.
- Levine, E., Spencer, J. L., Isard, S. A., Onstad, D. W. and Gray, M. E. (2002) Adaptation of the western corn rootworm to crop rotation: evolution of a new strain in response to a management practice. *American Entomologist* **48**: 94-107.
- Ludwick, D. C. and Hibbard, B. E. (2016) Rootworm management: status of GM traits, insecticides and potential new tools. *CAB Reviews* **11**: 1-10.
- Ludwick, D. C., Meihls, L. N., Ostlie, K. R., Potter, B. D., French, L. and Hibbard, B. E. (2017) Minnesota field population of western corn rootworm (Coleoptera: Chrysomelidae) shows incomplete resistance to Cry34Ab1/Cry35Ab1 and Cry3Bb1. *Journal of Applied Entomology* **141**: 28-40.

- Mathew, L. G., Ponnuraj, J., Mallappa, B., Chowdary, L. R., Zhang, J., Tay, W. T., *et al.* (2018) ABC transporter mis-splicing associated with resistance to Bt toxin Cry2Ab in laboratory- and field-selected pink bollworm. *Scientific Reports* **8**: 13531.
- Meihls, L. N., Higdon, M. L., Siegfried, B. D., Miller, N. J., Sappington, T. W., Ellersieck, M. R., *et al.* (2008) Increased survival of western corn rootworm on transgenic corn within three generations of on-plant greenhouse selection. *Proceedings of the National Academy of Sciences* **105**: 19177-19182.
- Melo, A. L. d. A., Soccol, V. T. and Soccol, C. R. (2016) *Bacillus thuringiensis*: mechanism of action, resistance, and new applications: a review. *Critical Reviews in Biotechnology* **36**: 317-326.
- Metzker, M. L. (2010) Sequencing technologies - the next generation. *Nature Review Genetics* **11**: 31-46.
- Miller, N. J., Estoup, A., Toepfer, S., Bourguet, D., Lapchin, L., Derridj, S., *et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**: 992.
- Miller, N. J., Guillemaud, T., Giordano, R., Siegfried, B. D., Gray, M. E., Meinke, L. J., *et al.* (2009) Genes, gene flow and adaptation of *Diabrotica virgifera virgifera*. *Agricultural and Forest Entomology* **11**: 47-60.
- Miller, N. J., Kim, K. S., Ratcliffe, S. T., Estoup, A., Bourguet, D. and Guillemaud, T. (2006) Absence of genetic divergence between western corn rootworms (Coleoptera: Chrysomelidae) resistant and susceptible to control by crop rotation. *Journal of Economic Entomology* **99**: 685-690.
- Miller, N. J., Richards, S. and Sappington, T. W. (2010) The prospects for sequencing the western corn rootworm genome. *Journal of Applied Entomology* **134**: 420-428.
- Miller, N. J. and Zhao, Z. (2015) Transcriptional responses of *Diabrotica virgifera virgifera* larvae to benzoxazinoids. *Journal of Applied Entomology* **139**: 416-423.
- Milner, R. J. (1994) History of *Bacillus thuringiensis*. *Agriculture, Ecosystems & Environment* **49**: 9-13.
- Ocelotl, J., Sánchez, J., Gómez, I., Tabashnik, B. E., Bravo, A. and Soberón, M. (2017) ABCC2 is associated with *Bacillus thuringiensis* Cry1Ac toxin oligomerization and membrane insertion in diamondback moth. *Scientific Reports* **7**: 2386.
- Okumura, S., Ohba, M., Mizuki, E., Crickmore, N., Côté, J.-C., Nagamatsu, Y., *et al.* (2010) Parasporin nomenclature. <http://parasporin.fitc.pref.fukuoka.jp/>
- Palma, L., Muñoz, D., Berry, C., Murillo, J. and Caballero, P. (2014) *Bacillus thuringiensis* toxins: an overview of their biocidal activity. *Toxins* **6**: 3296-3325.

- Pardo-López, L., Soberón, M. and Bravo, A. (2013) *Bacillus thuringiensis* insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection. *FEMS Microbiology Reviews* **37**: 3-22.
- Park, Y., Abdullah, M. A. F., Taylor, M. D., Rahman, K. and Adang, M. J. (2009) Enhancement of *Bacillus thuringiensis* Cry3Aa and Cry3Bb toxicities to Coleopteran larvae by a toxin-binding fragment of an insect cadherin. *Applied and Environmental Microbiology* **75**: 3086-3092.
- Portugal, L., Muñoz-Garay, C., Martínez de Castro, D. L., Soberón, M. and Bravo, A. (2017) Toxicity of Cry1A toxins from *Bacillus thuringiensis* to CF1 cells does not involve activation of adenylate cyclase/PKA signaling pathway. *Insect Biochemistry and Molecular Biology* **80**: 21-31.
- Pusztai, M., Fast, P., Gringorten, L., Kaplan, H., Lessard, T. and Carey, P. R. (1991) The mechanism of sunlight-mediated inactivation of *Bacillus thuringiensis* crystals. *Biochemical Journal* **273**: 43-47.
- Rault, L. C., Siegfried, B. D., Gassmann, A. J., Wang, H., Brewer, G. J. and Miller, N. J. (2018) Investigation of Cry3Bb1 resistance and intoxication in western corn rootworm by RNA sequencing. *Journal of Applied Entomology* **142**: 921-936.
- Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., Beeman, R. W., *et al.* (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**: 949.
- Rukmini, V., Reddy, C. Y. and Venkateswerlu, G. (2000) *Bacillus thuringiensis* crystal  $\delta$ -endotoxin: Role of proteases in the conversion of protoxin to toxin. *Biochimie* **82**: 109-116.
- Sanahuja, G., Banakar, R., Twyman, R. M., Capell, T. and Christou, P. (2011) *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnology Journal* **9**: 283-300.
- Sappington, T. W. (2011) Developing genomics tools for the western corn rootworm – progress and promise. *IWGO Newsletter* **31**: 8.
- Sappington, T. W., Siegfried, B. D. and Guillemau, T. (2006) Coordinated *Diabrotica* genetics research: accelerating progress on an urgent insect pest problem. *American Entomologist* **52**: 90-97.
- Sato, R., Adegawa, S., Li, X., Tanaka, S. and Endo, H. (2019) Function and role of ATP-Binding cassette transporters as receptors for 3D-Cry toxins. *Toxins* **11**: 124.
- Schnepf, H. E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J. A., Feitelson, J., *et al.* (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews* **62**: 775-806.

- Schnepf, H. E. and Whiteley, H. R. (1981) Cloning and expression of the *Bacillus thuringiensis* crystal protein gene in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **78**: 2893-2897.
- Schoville, S. D., Chen, Y. H., Andersson, M. N., Benoit, J. B., Bhandari, A., Bowsher, J. H., *et al.* (2018) A model species for agricultural pest genomics: the genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Scientific Reports* **8**: 1931.
- Schrader, P. M., Estes, R. E., Tinsley, N. A., Gassmann, A. J. and Gray, M. E. (2017) Evaluation of adult emergence and larval root injury for Cry3Bb1-resistant populations of the western corn rootworm. *Journal of Applied Entomology* **141**: 41-52.
- Segura-León, O. L. (2004) Phylogeography of *Diabrotica virgifera* LeConte and *Diabrotica virgifera zea* Krysan and Smith (Coleoptera: Chrysomelidae). Doctor of Philosophy. University of Nebraska-Lincoln.
- Sha, J., Zhang, J., Chi, B., Liu, R., Li, H. and Gao, J. (2018) Sip1Ab gene from a native *Bacillus thuringiensis* strain QZL38 and its insecticidal activity against *Colaphellus bowringi* Baly. *Biocontrol Science and Technology* **28**: 459-467.
- Siegfried, B. D., Waterfield, N. and French-Constant, R. H. (2005) Expressed sequence tags from *Diabrotica virgifera virgifera* midgut identify a coleopteran cadherin and a diversity of cathepsins. *Insect molecular biology* **14**: 137-143.
- Smouse, D. and Nishiura, J. (1997) A *Bacillus thuringiensis*  $\delta$ -endotoxin induces programmed cell death in mosquito larvae. *Cell Death And Differentiation* **4**: 560.
- Soberón, M., Pardo-López, L., López, I., Gómez, I., Tabashnik, B. E. and Bravo, A. (2007) Engineering modified Bt toxins to counter insect resistance. *Science* **318**: 1640-1642.
- Tabashnik, B. E. and Carrière, Y. (2017) Surge in insect resistance to transgenic crops and prospects for sustainability. *Nature Biotechnology* **35**: 926.
- Tanaka, S., Endo, H., Adegawa, S., Iizuka, A., Imamura, K., Kikuta, S., *et al.* (2017) *Bombyx mori* ABC transporter C2 structures responsible for the receptor function of *Bacillus thuringiensis* Cry1Aa toxin. *Insect Biochemistry and Molecular Biology* **91**: 44-54.
- Tanaka, S., Endo, H., Adegawa, S., Kikuta, S. and Sato, R. (2016) Functional characterization of *Bacillus thuringiensis* Cry toxin receptors explains resistance in insects. *The FEBS Journal* **283**: 4474-4490.
- Tanaka, S., Miyamoto, K., Noda, H., Jurat-Fuentes, J. L., Yoshizawa, Y., Endo, H., *et al.* (2013) The ATP-binding cassette transporter subfamily C member 2 in *Bombyx mori* larvae is a functional receptor for Cry toxins from *Bacillus thuringiensis*. *The FEBS Journal* **280**: 1782-1794.

Tanaka, S., Yoshizawa, Y. and Sato, R. (2012) Response of midgut epithelial cells to Cry1Aa is toxin-dependent and depends on the interplay between toxic action and the host apoptotic response. *The FEBS Journal* **279**: 1071-1079.

Vachon, V., Laprade, R. and Schwartz, J.-L. (2012) Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: A critical review. *Journal of Invertebrate Pathology* **111**: 1-12.

Walters, F. S., deFontes, C. M., Hart, H., Warren, G. W. and Chen, J. S. (2010) Lepidopteran-active variable-region sequence imparts coleopteran activity in eCry3.1Ab, an engineered *Bacillus thuringiensis* hybrid insecticidal protein. *Applied and Environmental Microbiology* **76**: 3082-3088.

Wang, H., Coates, B. S., Chen, H., Sappington, T. W., Guillemaud, T. and Siegfried, B. D. (2013) Role of a gamma-aminobutyric acid (GABA) receptor mutation in the evolution and spread of *Diabrotica virgifera virgifera* resistance to cyclodiene insecticides. *Insect Molecular Biology* **22**: 473-484.

Wang, H., Eyun, S.-i., Arora, K., Tan, S., Gandra, P., Moriyama, E., *et al.* (2017) Patterns of gene expression in western corn rootworm (*Diabrotica virgifera virgifera*) neonates, challenged with Cry34Ab1, Cry35Ab1 and Cry34/35Ab1, based on next-generation sequencing. *Toxins* **9**: 124.

Wangila, D. S., Gassmann, A. J., Petzold-Maxwell, J. L., French, B. W. and Meinke, L. J. (2015) Susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to Bt corn events. *Journal of Economic Entomology* **108**: 742-751.

Xiao, Y. and Wu, K. (2019) Recent progress on the interaction between insects and *Bacillus thuringiensis* crops. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**: 20180316.

Xiao, Y., Zhang, T., Liu, C., Heckel, D. G., Li, X., Tabashnik, B. E., *et al.* (2014) Mis-splicing of the ABCC2 gene linked with Bt toxin resistance in *Helicoverpa armigera*. *Scientific Reports* **4**: 6184.

Xu, C., Wang, B.-C., Yu, Z. and Sun, M. (2014) Structural insights into *Bacillus thuringiensis* Cry, Cyt and parasporin toxins. *Toxins* **6**: 2732-2770.

Zhang, F., Peng, D., Cheng, C., Zhou, W., Ju, S., Wan, D., *et al.* (2016) *Bacillus thuringiensis* Crystal Protein Cry6Aa Triggers *Caenorhabditis elegans* Necrosis Pathway Mediated by Aspartic Protease (ASP-1). *PLOS Pathogens* **12**: e1005389.

Zhang, T., Coates, B. S., Wang, Y., Wang, Y., Bai, S., Wang, Z., *et al.* (2017) Down-regulation of aminopeptidase N and ABC transporter subfamily G transcripts in Cry1Ab and Cry1Ac resistant Asian corn borer, *Ostrinia furnacalis* (Lepidoptera: Crambidae). *International journal of biological sciences* **13**: 835-851.

- Zhang, X., Candas, M., Griko, N. B., Rose-Young, L. and Bulla, L. A. (2005) Cytotoxicity of *Bacillus thuringiensis* Cry1Ab toxin depends on specific binding of the toxin to the cadherin receptor BT-R1 expressed in insect cells. *Cell Death And Differentiation* **12**: 1407.
- Zhang, X., Candas, M., Griko, N. B., Taussig, R. and Bulla, L. A. (2006) A mechanism of cell death involving an adenylyl cyclase/PKA signaling pathway is induced by the Cry1Ab toxin of *Bacillus thuringiensis*. *Proceedings of the National Academy of Sciences* **103**: 9897-9902.
- Zhang, X., Griko, N. B., Corona, S. K. and Bulla, L. A. (2008) Enhanced exocytosis of the receptor BT-R1 induced by the Cry1Ab toxin of *Bacillus thuringiensis* directly correlates to the execution of cell death. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **149**: 581-588.
- Zhao, Z., Meihls, L. N., Hibbard, B. E., Ji, T., Elsik, C. G. and Shelby, K. S. (2019) Differential gene expression in response to eCry3.1Ab ingestion in an unselected and eCry3.1Ab-selected western corn rootworm (*Diabrotica virgifera virgifera* LeConte) population. *Scientific Reports* **9**: 4896.
- Zukoff, S. N., Ostlie, K. R., Potter, B., Meihls, L. N., Zukoff, A. L., French, L., *et al.* (2016) Multiple assays indicate varying levels of cross resistance in Cry3Bb1-selected field populations of the western corn rootworm to mCry3A, eCry3.1Ab, and Cry34/35Ab1. *Journal of Economic Entomology* **109**: 1387-1398.

## **Chapter 2**

Differential gene expression in response to eCry3.1Ab ingestion in an unselected and eCry3.1Ab-selected western corn rootworm (*Diabrotica virgifera virgifera* LeConte) population

## Abstract

The western corn rootworm (WCR, *Diabrotica virgifera virgifera* LeConte) is one of the most destructive pests in the U.S. Corn Belt. Transgenic maize lines expressing various Cry toxins from *Bacillus thuringiensis* have been adopted as a management strategy. However, resistance to many Bt toxins has occurred. To investigate the mechanisms of Bt resistance we carried out RNA-seq using Illumina sequencing technology on laboratory-selected eCry3.1Ab-resistant and susceptible WCR neonates, which fed on seedling maize with and without eCry3.1Ab for 12 and 24 hours. In a parallel experiment, RNA-seq experiments were conducted on the midgut tissue of neonate WCR from the same treatments. After *de novo* transcriptome assembly, we identified differentially expressed genes (DEGs). Results from the assemblies and annotation indicate that WCR neonates from the eCry3.1Ab-selected resistant colony expressed a small number of up and down-regulated genes following Bt intoxication. In contrast, unselected susceptible WCR neonates expressed a large number of up- and down-regulated transcripts in response to intoxication. Annotation and pathway analysis of DEGs between susceptible and resistant whole WCR and their midgut tissue revealed genes associated with cell membrane, immune response, detoxification, and potential Bt receptors which are likely related to eCry3.1Ab resistance. This research provides a framework to study the toxicology of Bt toxins and mechanism of resistance in WCR, an economically important Coleopteran pest species.

## Introduction

The *Bacillus thuringiensis* (Bt) Cry3  $\delta$ -endotoxin superfamily is known for its specificity to Coleopteran species (Lee et al. 2013; Perlak et al. 1993). Transgenic maize hybrids expressing Cry3Bb1, mCry3A, eCry3.1Ab, as well as Cry34/35Ab1 have been introduced to the market to manage western corn rootworm (WCR, *Diabrotica virgifera virgifera* LeConte). As a highly adaptive species, WCR has developed resistance to broadcast soil insecticides (Ball and Weekman 1962), aerial spray insecticides for reducing adult number (Meinke et al. 1998; Pereira et al. 2015), crop rotation (Levine et al. 2002), and Cry3 Bt toxins. Laboratory selection experiments have developed WCR colonies resistant to Cry3Bb1, mCry3A, Cry34/35Ab1 and eCry3.1Ab by continuously rearing WCR on transgenic maize lines (Frank et al. 2013; Lefko et al. 2008; Meihls et al. 2011; Meihls et al. 2008). Resistance to Cry3  $\delta$ -endotoxin has become a practical issue since field resistant WCR populations have been reported in many locations (Gassmann et al. 2011; Ludwick et al. 2017; Wangila et al. 2015; Zukoff et al. 2016). Although eCry3.1Ab is the most recent Bt toxin (Walters et al. 2010) to the market without reported control failure in the field, laboratory selection (Frank et al. 2013) and cross-resistance experiments (Zukoff et al. 2016) indicate that resistance to eCry3.1Ab is likely in the field due to the cross-resistance to other Bt toxins (Ludwick and Hibbard 2016).

Three Cry toxins (Cry3Bb1, mCry3A and eCry3.1Ab) for WCR control are derived from the Cry3 superfamily. In Lepidopteran insects, the consumption of Bt protoxin by larvae is followed by activation via cleavage by midgut proteases, binding to brush border receptors, toxin insertion into the epithelial cell membrane, pore formation

by oligomerization, and finally, destruction of midgut epithelial cells (Bravo et al. 2007; Knowles 1994) followed by fatal sepsis (Caccia et al. 2016). Anything interfering with these processes (e.g. decreased activation, reduced receptor affinity, increased toxin degradation, or increased repair of midgut epithelial cells) could cause resistance (Heckel et al. 2007). However, the detailed interactions between Cry3 toxins and WCR is less known. Similar to Lepidopteran insects, reduced binding of mCry3A to the midgut epithelial cell membrane has been observed in resistant WCR colonies (Zhao et al. 2016). However, the Cry3 receptor has not been validated in WCR. Although potential receptors including cadherin-like protein (Sayed et al. 2007) and ATP-binding cassette (ABC) transporters (Flagel et al. 2015) are present, their roles in Cry3 mode of action are remained unclear.

Toxin dose is a very important component of what has been coined the “high-dose/refuge strategy” to delay resistance (Ludwick et al. 2017). In this theory the “high-dose” refers to Bt toxins that cause very high mortality to pest population. The initial frequency of alleles conferring resistance is low, fitness costs of resistance are present, mating between resistant and susceptible insects is random, and resistance to the Bt toxin is recessive (Gould 1998; Tabashnik and Gould 2012). For Lepidopteran insects e.g. *Heliothis virescens* (Fabricius), the Cry1Ac is considered as a high-dose toxin, which is able to kill the insects with susceptible alleles. Only individuals with homozygous resistant alleles will survive, yielding a population with extremely high resistance ratio to Bt toxin. The resistant genes could be identified by QTL mapping (Gahan et al. 2001; Gahan et al. 2010) from homozygous resistant population. In WCR, the situation is less optimal. First, all four WCR-targeting Bt toxins are considered as low to moderate dose

(Clark et al. 2012; Hibbard et al. 2010a; Hibbard et al. 2010b; Storer et al. 2006), where individuals with heterozygous resistance alleles may survive. Second, there might be more than one gene involved in Bt resistance in WCR. The resistant genes selected in laboratory-selected resistant populations may not be sufficient to estimate the resistant alleles in field resistant populations (French-Constant et al. 2004). Third, the lack of a well-assembled, high-quality WCR genome sequence hampers the application of advanced genetic and molecular approaches to analyze resistant-related variation at the genomic level. RNA-seq followed by de novo assembly is an alternative for analyzing gene expression without a reference genome. In WCR, RNA-seq has been successfully used to study callus digestion (Eyun et al. 2014), interaction with corn defense chemicals (Miller and Zhao 2015), adaptation to pesticides (Coates et al. 2016), resistance to crop rotation (Chu et al. 2015) and response to Bt toxins e.g. Cry34/35Ab1 (Wang et al. 2017), and Cry3Bbb1 (Rault et al. 2018).

To investigate the gene response to eCry3.1Ab intoxication and mechanisms of resistance, we sequenced the transcriptome of eCry3.1Ab-selected resistant and unselected, susceptible WCR feeding with maize root with and without eCry3.1Ab for 12 and 24 hours. Moreover, to further track down the tissue interactions with Bt toxin we sequenced the transcriptome of midgut from recovered neonates with the same treatment. The expression patterns of different populations under these conditions reveal how WCR respond to eCry3.1Ab and provide clues of resistance mechanisms.

## **Material and Methods**

### **Insects and bioassay**

The eCry3.1Ab-resistant WCR colony was initially selected and reared on non-elite non-commercial eCry3.1Ab-expressing transgenic maize (event 5130) under laboratory conditions (Frank et al. 2013). Both the eCry3.1Ab-resistant and unselected susceptible colonies were developed from a single population. The eCry3.1Ab-resistant and susceptible colonies had been maintained on eCry3.1Ab-expressing transgenic maize (material ID 12MG00345) and its near-isoline (material ID 12MG001181) respectively for more than 30 generations. For the current experimental design (Table 2-1), both eCry3.1Ab-expressing and isoline maize seeds were surface sterilized and germinated in Petri dishes with moistened filter paper at 23°C for 4-6 days without illumination. Approximately 30 neonates hatched within 24 hours were transferred to a Petri dish containing 3-4 seedlings of each line. After 12 or 24 hours feeding, the living first-instar larvae were recovered. In a separate identical experiment, the midgut was dissected from 30-40 recovered larvae of each Petri dish. Both whole larvae and midgut were flash frozen in liquid nitrogen and stored at -80°C until processing. Both the whole larvae and midgut bioassays were replicated independently three times as full biological replications.

### **RNA extraction, library construction and sequencing**

RNA was extracted using Trizol and purified by Direct-zol RNA Mini Prep kit (Zymo Research, Irving, CA). DNase treatment (ThermoFisher, Waltham, MA) was incorporated to remove genomic DNA contamination. RNA samples were checked for integrity using a fragment analyzer (University of Missouri DNA Core Laboratory). Strand-specific RNA-Seq libraries were prepared using the Illumina TruSeq HT Stranded Total RNA Library Prep Kit (Illumina, San Diego, CA), following the manufacturer's

instructions. For whole larvae and midgut respectively, 24 libraries (2 insect colonies  $\times$  2 corn lines  $\times$  2 times  $\times$  3 biological replications) were normalized, pooled and sequenced on two lanes of Illumina HiSeq 2000 sequencer using 100-nucleotide pair-end protocol (Global Biologics LLC, Columbia, MO, USA).

### **De novo assembly of transcriptome**

Adapters and low quality reads were trimmed using FastqMcf (version 1.04.803, <https://github.com/ExpressionAnalysis/ea-utils/blob/wiki/FastqMcf.md>) and Trimmomatic (version 0.36) (Bolger et al. 2014), respectively. To remove the mtDNA, the reads were aligned to the WCR mitochondrial genome (Coates 2014) using Bowtie 2 (version 4.7.7) (Langmead and Salzberg 2012) with default settings. Paired reads concordantly aligned with no mismatch were considered as mitochondrial reads and discarded. These steps resulted in “clean reads” for assembly and differential expression analysis.

Trinity (version r2013-11-10) (Grabherr et al. 2011) was used to *de novo* assemble larval and midgut transcriptomes using cleaned reads pooled from libraries of each sample type with default setting. The unigene set of each transcriptome was obtained by removing sequence with over 95% of similarity in Blast2GO (version 4.1.9) (Conesa et al. 2005). BUSCO analyses (version v3) with the lineage “insect\_obd9” (Simão et al. 2015) were performed on transcriptome assemblies and unigene sets to evaluate their quality and completeness.

### **Differential expression analysis**

Cleaned reads were aligned to the corresponding transcriptome unigene set using Bowtie 2 with default pair-end settings (Langmead and Salzberg 2012). The output SAM

files were converted to BAM format using SAMtools (version 0.1.20) (Li et al. 2009). The differential expression analysis was conducted in R (version 3.4.1) (Ihaka and Gentleman 1996). The read counts were loaded in R using the GenomicAlignment (version 1.6.3) and GenomicRanges (version 1.22.4) packages (Lawrence et al. 2013). We counted only concordant alignment pairs while accepting multiple mapping reads due to the potential existence of isoforms in the unigene sets. The low expression unigenes were removed by applying filters with at least 2 count per million (CPM) over 3 samples. The differentially expressed contigs were assessed using the edgeR-robust algorithm of the edgeR package (version 3.12.1) (McCarthy et al. 2012; Robinson et al. 2010) with the trimmed mean of M-values (TMM) normalization method (Robinson and Oshlack 2010; Zhou et al. 2014). False discovery rate (FDR) was controlled at 0.05 by the edgeR package and was used to determine the significance of differentially expressed genes (DEGs).

### **Annotation and pathway analysis**

Blast2GO was used for gene annotation, enrichment and pathway analysis. Transcriptomes were annotated by BLASTX against NCBI non-redundant (NR) database using an E-value cutoff of 1.0E-3. InterProScan was used to identify protein domains from 11 member databases (Quevillon et al. 2005). Gene ontology terms (GO) were assigned based on the results of BLASTX and InterProScan. Gene set enrichment analysis (GSEA) was used to determine enriched GO terms by comparing the gene list versus the entire transcriptome. The enzyme codes were assigned to predicted enzymes function based on their GO terms. Those enzymes were mapped to Kyoto Encyclopedia

of Gene and Genomes (KEGG) database for pathway analysis (Kanehisa and Susumu 2000).

## **Results and Discussion**

### **Transcriptome assembly and annotation of WCR transcriptome**

The transcriptome of whole larvae and midgut were separately *de novo* assembled. Reads from either whole larvae or midgut were individually pooled to increase the assembly coverage. Low quality and mitochondrial sequences were removed prior to assembly. Transcriptome assembly resulted in a whole larval transcriptome with 204,842 contigs from 57 Gb of reads and midgut transcriptome with 226,115 contigs from 137 Gb of reads. The two transcriptomes had comparable average sequence length (measured as N50, a weighted median statistic that more than half of the nucleotides of a transcriptome belong to the contigs of this size N50 or longer), GC content and length distribution (Table 2-2, Figure 2-1). After removing duplicate contigs with more than 95% sequence similarity, we obtained 187,570 and 209,167 contigs respectively from larval and midgut transcriptome (Table 2-2). Herein, we refer to the two reduced redundancy sets of contigs as “unigene” sets. BUSCO analyses (Simão et al. 2015) were applied to evaluate the completeness of each unigene set. All of the transcriptome and unigene sets cover over 95% of Insecta single-copy orthologs. Both unigene sets maintain the coverage and integrity while duplication is reduced, indicating the acceptable qualities for functional analysis (Table 2-3).

Transcriptomes and their unigene sets were aligned to the NCBI non-redundant protein database (NR) using BLASTX. Only 42% of larval and 38.2% of midgut unigenes had significant BLASTX results. The species distribution of BLASTX top hits

indicated that a predominant number of unigenes could be annotated by the Coleopteran model species *Tribolium castaneum* (Herbst) (Table 2-4, Figure 2-2). We also identified 372 and 400 unigenes from larval and midgut transcriptome, respectively, that were directly aligned to genes from *Diabrotica* species (Table 2-4).

### **Differential expression analysis of WCR transcriptome**

To identify the genes involved in eCry3.1Ab response, we analyzed the differentially expressed unigenes from whole larval and midgut transcriptomes between eCry3.1Ab-feeding and non-Bt isoline feeding WCR at both 12- and 24-hour time points. To understand the expression differences between eCry3.1Ab-selected and susceptible, control WCR from the same original population, we also compared the expression differences in whole larvae and midgut between two populations when being fed with the same kind of maize root. After the alignment and filtering out unigenes with extremely low expression levels, only 31,875 of larval and 22,954 of midgut transcriptome unigenes were proceeded to edgeR analysis (Robinson et al. 2010). The patterns of differentially expressed genes are shown (Figure 2-3). Regardless of exposure time or tissue type, susceptible WCR differentially expressed a much larger number of genes in response to eCry3.1Ab intoxication. In contrast, eCry3.1-selected WCR had dramatically fewer differentially expressed genes (DEGs). The selected and unselected WCR share many DEGs in both whole larvae and midgut transcriptome, while some unique DEGs were colony specific, especially in unselected colony (Figure 2-4).

Albeit smaller in size, we identified DEGs between eCry3.1Ab-selected and unselected WCR, especially when both groups of neonates were fed on eCry3.1Ab maize root. The function of those DEGs and the pathways in which they were involved may

reveal the physiological differences and the mechanism of eCry3.1Ab resistance in the selected population.

### **GO and pathway analysis on eCry3.1Ab feeding WCR midgut**

To further investigate the molecular and physiological adaptation to intoxication on eCry3.1Ab-selected WCR we applied gene ontology (GO) analysis to differentially expressed genes using Blast2GO. We compared the top 20 level-2 GO terms called from the unigene set from midgut DEGs between selected and unselected WCR feeding on eCry3.1Ab maize for 24 h. The GO terms were ranked based on number of unigenes of each GO term. The terms “metabolic process”, “catalytic activity” and “membrane” accounted for the majority of each of the three ontologies (Biological Process (BP), Molecular Function (MF) and Cellular Component (CC), respectively), suggesting the primary location and functions of eCry3.1Ab resistance. Other than that, the BP term “cellular process” and MF term “binding” also implied the binding and processing of toxin may have a role in eCry3.1Ab sensitivity.

We compared the distribution of level-2 GO terms between the DEG unigene set described above and midgut transcriptome unigene set (Table 2-5). The BP terms “Multiple-organism process”, “immune system process”, “detoxification” and “cell killing” and the CC terms “extracellular region” and “extracellular part” were among the most over represented GOs in DEG unigene sets, while the BP term “reproduction” and MF term “structural molecule activity” were the most under represented. However gene set enrichment analysis (GSEA) showed that no specific GO term was significantly enriched in the DEG unigene set.

We further predicted enzyme codes (EC) from GO terms for unigenes and used the EC to map differentially expressed unigenes to pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Susumu 2000). The differentially expressed midgut unigenes were involved in 49 pathways, including purine metabolism (KEGG ID: 00230), glutathione metabolism (KEGG ID: 00480), fatty acid synthesis (KEGG ID: 00061), glycerophospholipid metabolism (KEGG ID: 00564) and drug metabolism (KEGG ID: 00983). The fatty acid synthesis pathway was also identified in whole larval differentially expressed unigenes. These results suggest that differences in eCry3.1Ab tolerances might arise from the alternation of genes related to detoxification, membrane functions and metabolism.

### **Expression of potential and novel eCry3.1Ab resistant genes**

Research has revealed that cadherins (Gahan et al. 2001), ABC transporters (Gahan et al. 2010), and aminopeptidase N (APN) (Bravo et al. 2004) are Bt receptors in lepidopteran insects. Both cadherins and ABC transporters have been identified in WCR (Flagel et al. 2015; Sayed et al. 2007). Upon BLASTX, we found 102 unigenes with cadherin function, 209 unigenes with ABC transporter function, and 50 unigenes with APN function from the midgut unigene set. In whole larval unigene set, there were 78 unigenes with cadherin function, 191 unigenes with ABC transporter function, and 62 unigenes with APN function.

The midgut gene expression profile showed that cadherin was not significantly differentially expressed, while some ABC transporter subfamily C (a.k.a. multidrug resistance-associated protein) and APN genes were differentially expressed following Cry3.1Ab ingestion, especially in susceptible WCR (Figure 2-5). The expression level of

one APN midgut unigene (comp127589\_c1\_seq1) was significantly higher in susceptible WCR midgut, when both selected and susceptible-WCR were given eCry3.1Ab for 24 hours. Two ABC transporter unigenes (comp121889\_c0\_seq1, comp126268\_c0\_seq1) showed the same pattern, but the increased expression in susceptible WCR was not significant. The protein structure of eCry3.1Ab toxin contains partial Cry1Ab sequence at its C-terminus (Walters et al. 2010). In many Lepidopteran insects, APNs serve as binding receptors of Cry1A toxins. The mutations or reduced expression of APN resulted in resistance to Bt toxins (Bravo et al. 2004). Our results suggest that APN is a potential eCry3.1Ab target and the reduced expression level under intoxication might contribute to resistance in WCR.

Proteases have been associated with Bt resistance either by increasing digestion of toxins (Loseva et al. 2002; Rausell et al. 2007), or by decreasing the proteolytic activation of Bt pro-toxins (Wei et al. 2016). We identified 112 unigenes with metalloprotease functions and 199 unigenes with cathepsin functions in the larval unigene set. In the midgut unigene set, there were 131 unigenes with metalloproteases functions and 234 unigenes with cathepsin functions. Expression levels of some proteases were regulated by Bt intoxication either in selected or susceptible WCR, while none of those digestive proteases was up regulated in selected WCR midgut compared to the susceptible one when both insects were fed with eCry3.1Ab maize.

Considering the expression patterns, GO annotation, and pathway analysis, we infer that at least two novel genes are likely involved in the resistance to eCry3.1Ab. Esterase has been reported involved in Cry1Ac resistance in *Helicoverpa armigera* (Gunning et al. 2005). In WCR, we observed a novel esterase (comp36305\_c0\_seq1)

expressed in the midgut of selected colony was higher than unselected one after 24 hours of intoxication. The second gene is a dynein heavy chain-like protein (comp127369\_c3\_seq14). It was constitutively up regulated in selected WCR regardless of diet and time. Since dynein is a cytoskeletal motor protein involved in intracellular transportation and the movement of chromosomes, we propose that selected-WCR may have a stronger activity in either endocytosis or cell mitosis to remove the attached eCry3.1Ab molecules (Bonfini et al. 2016), or to repair damaged epithelial cells (Castagnola and Jurat-Fuentes 2016).

### **Conclusion**

After a comprehensive analysis of an RNA-seq experiment with eCry3.1Ab-selected and susceptible, control WCR, the transcriptome, unigene sets and reads provided numerous resources for studying the interaction between Cry3 and this coleopteran species. This study is the first step approaching the delineation of Bt resistance mechanisms in WCR. We propose more than one potential mechanism of resistance to eCry3.1Ab – a dual action Bt toxin. With the recent published WCR genome sequence, future research will detect the genomic-wide genetic variations associated with Bt resistance. Studies to explore how Bt toxins affect gene alternative splicing and whether the alternative sliced genes are related to Bt resistance in WCR. We are also developing cellular and molecular methods i.e. cultured cell expression, RNAi gene silencing, individual genotyping to further study the detail mechanism of resistance to eCry3.1Ab as well as other Bt toxins. The continuous discoveries in this field will lead to improving strategies for insect resistant management and the developing of novel entomotoxins.

## **Acknowledgments**

We appreciate Julie Barry (USDA-ARS Plant Genetics Research Unit) for maintaining and providing WCR colonies for this research. Michelle Gregory (USDA-ARS- BCIRL) assisted with some of the experiments. Syngenta Biotechnology provided the transgenic and near-isoline maize seeds. Colin Diesh, Deepak Unni and Justin Le Tourneau from the Division of Animal Sciences, University of Missouri maintained the server for transcriptome data analysis and assisted with the bioinformatics software.

**Table 2-1.** The design of feeding assay of eCry3.1Ab-resistant and susceptible WCR colony. The same experiments were conducted independently three times each of whole larvae and midgut RNA-seq.

<b>Code</b>	<b>WCR colony</b>	<b>Maize Seedling Type</b>	<b>Feeding period (hrs)</b>
ERB12	eCry3.1Ab-resistant	eCry3.1Ab	12
ERB24	eCry3.1Ab-resistant	eCry3.1Ab	24
ERI12	eCry3.1Ab-resistant	Non-Bt Isoline	12
ERI24	eCry3.1Ab-resistant	Non-Bt Isoline	24
ESB12	Susceptible	eCry3.1Ab	12
ESB24	Susceptible	eCry3.1Ab	24
ESI12	Susceptible	Non-Bt Isoline	12
ESI24	Susceptible	Non-Bt Isoline	24

**Table 2-2.** Summary statistics of WCR whole larvae and midgut transcriptome assemblies and their unigene sets.

Whole Larval Transcriptome		Midgut Transcriptome		
<b>Input Reads</b>				
Raw reads pair	313,684,316		750,660,524	
Filtered reads pair	284,268,606		727,149,108	
Filtered total bases	57,171,701,994		137,820,263,438	
<b>Transcriptome Statistics</b>				
	<b>Trinity Assembly</b>	<b>Unigene Set</b>	<b>Trinity Assembly</b>	<b>Unigene Set</b>
Total assembled bases	173,452,950	130,813,521	176,268,219	136,641,964
Number of contigs	204,842	187,570	226,115	209,167
Average contig length	847	697.4117	780	653.2673
Min contig length	201	201	201	201
Max contig length	31,383	31,119	28,475	27,353
Number of contigs > 1kb	48,389	34,466	46,762	33,637
Number of contigs > 5kb	2,584	1,014	2,346	946
Number of contigs > 10kb	216	59	191	56
N50	1,523	1,096	1,351	977
GC content (%)	35.91	35.45	36.54	36.36

**Table 2-3.** Summary of BUSCO analysis of WCR whole larval and midgut transcriptome assemblies and their unigene sets.

	Whole Larvae		Midgut	
	Trinity Assembly	Unigene Set	Trinity Assembly	Unigene Set
Complete (%)	95.7	95.5	96.6	96.6
Duplicated (%)	27.6	21.7	31.2	23.7
Fragment (%)	3.3	3.5	1.7	1.7
Missing (%)	1.0	1.0	1.7	1.7

**Table 2-4.** Summary of BLASTX of WCR whole larval and midgut transcriptome assemblies and their unigene sets.

	Whole Larvae		Midgut	
	Trinity Assembly	Unigene set	Trinity Assembly	Unigene set
Total Contigs for BLASTX	204,842	187,570	226,115	209,167
Contigs with BLASTX Hits	86,123	73,699	86,411	75,388
Contigs with Coleopteran Hits				
<i>Diabrotica</i> spp.	292	372	315	400
<i>Tribolium castaneum</i>	25,617	34,936	22,262	33,924

**Table 2-5.** The top level-2 GO terms of DEG unigene set and midgut transcriptome unigene sets. The DEGs are from comparison of midgut gene expression between select and unselected-WCR when both were given eCry3.1Ab-expressing roots for 24 hours. The GO terms are ranked by number of sequences of each GO, as well as the percentage of sequences among each unigene set.

**A. Biological Process (BP) terms**

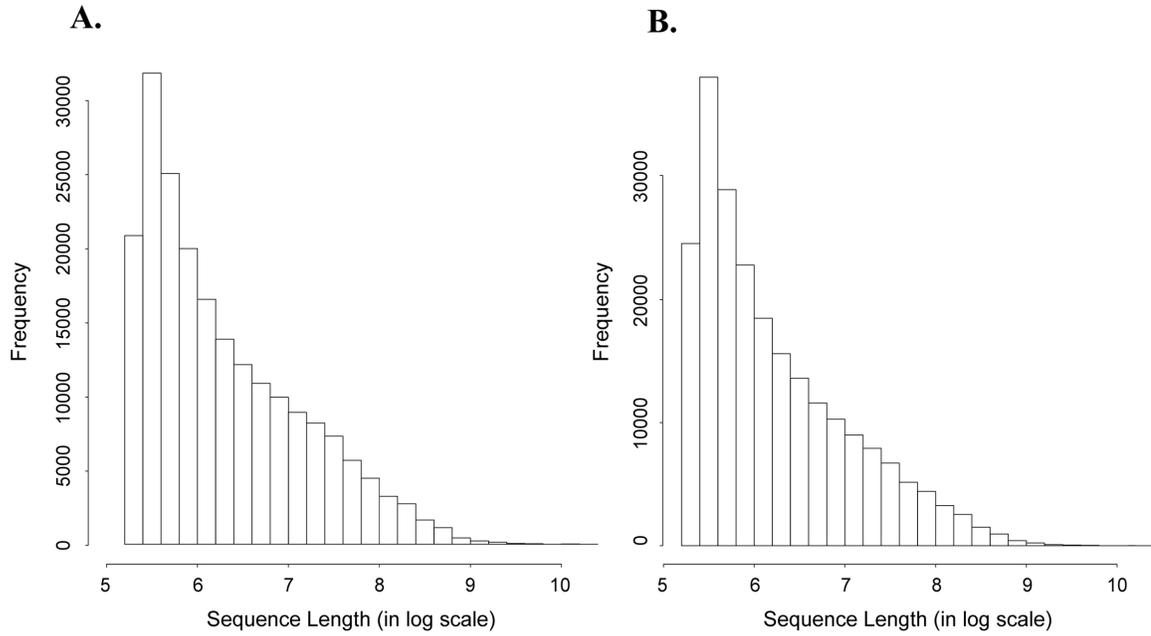
GO Terms	DEG Unigene set		Midgut Unigene Set	
	Percentage	Rank	Percentage	Rank
Metabolic process	29.52%	1	8.56%	1
Single-organism process	15.05%	2	3.78%	3
Cellular process	13.60%	3	8.09%	2
Localization	7.09%	4	1.58%	6
Biological regulation	2.75%	5	1.98%	4
Response to stimulus	2.17%	6	1.32%	7
Regulation of biological process	2.03%	7	1.79%	5
Multicellular organismal process	1.45%	8	0.54%	10
Developmental process	1.16%	9	0.47%	11
Cellular component organization or biogenesis	0.87%	10	1.01%	8
Signaling	0.87%	10	0.92%	9
Negative regulation of biological process	0.72%	12	0.21%	13
Positive regulation of biological process	0.58%	13	0.21%	12
Immune system process	0.58%	13	0.11%	18
Multi-organism process	0.43%	15	0.14%	16
Detoxification	0.29%	16	0.03%	22
Cell killing	0.29%	16	0.01%	24
Reproductive process	0.14%	18	0.14%	15
Reproduction	0.14%	19	0.14%	14
Locomotion	0.14%	19	0.09%	19

## B. Molecular Function (MF) terms

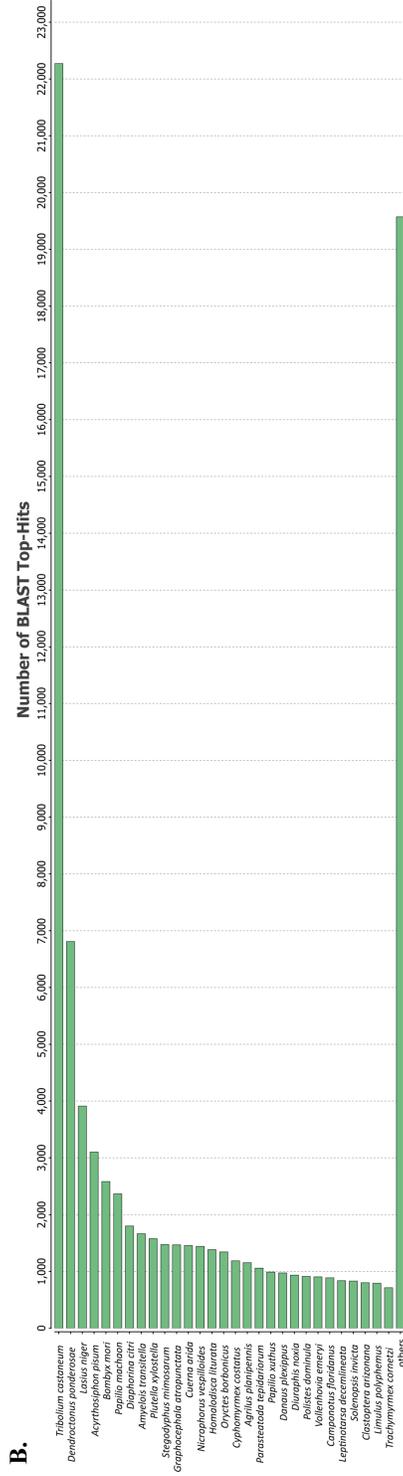
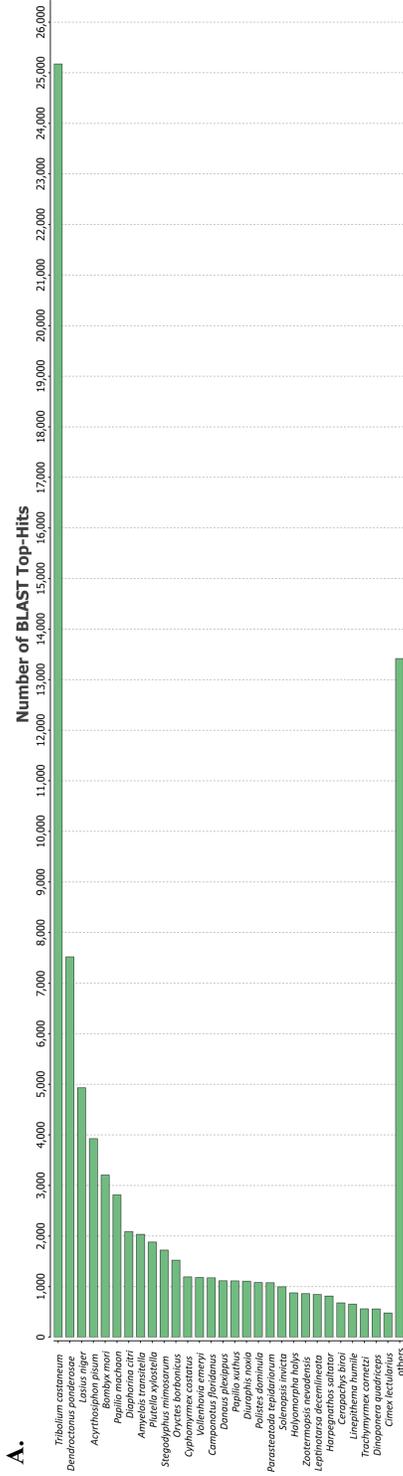
GO Terms	DEG Unigene set		Midgut Unigene Set	
	Percentage	Rank	Percentage	Rank
Catalytic activity	32.85%	1	8.52%	2
Binding	20.98%	2	8.79%	1
Transporter activity	3.62%	3	0.83%	3
Molecular transducer activity	0.72%	4	0.27%	6
Molecular function regulator	0.43%	5	0.23%	7
Structural molecule activity	0.43%	5	0.82%	4
Electron carrier activity	0.43%	5	0.08%	9
Signal transducer activity	0.29%	8	0.30%	5
Antioxidant activity	0.29%	8	0.05%	10

## C. Cellular Component (CC) terms

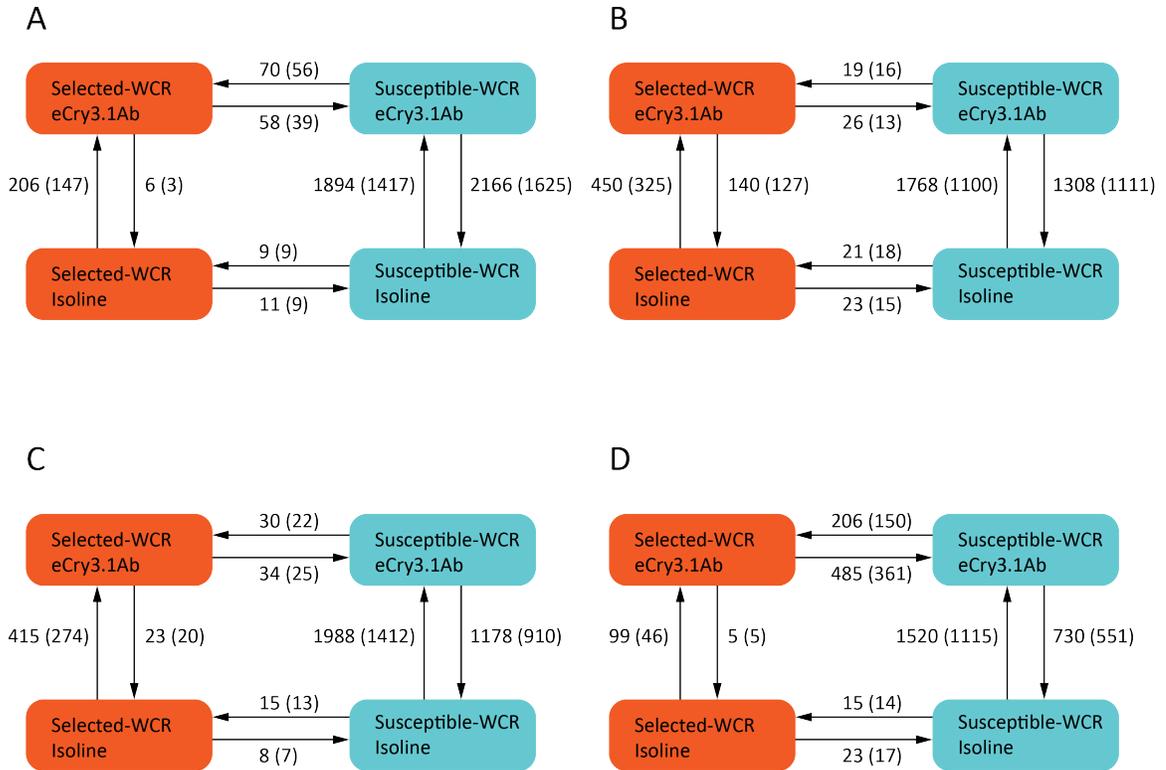
GO Terms	DEG Unigene set		Midgut Unigene Set	
	Percentage	Rank	Percentage	Rank
Membrane	9.12%	1	3.64%	3
Cell	5.93%	2	4.14%	1
Cell part	5.79%	3	4.09%	2
Extracellular region	5.35%	4	0.25%	9
Membrane part	5.07%	5	2.43%	5
Organelle	4.05%	6	2.69%	4
Extracellular region part	2.46%	7	0.09%	11
Organelle part	1.88%	8	1.43%	7
Macromolecular complex	0.87%	9	1.74%	6
Supramolecular complex	0.43%	10	0.10%	10
Membrane-enclosed lumen	0.43%	10	0.37%	8



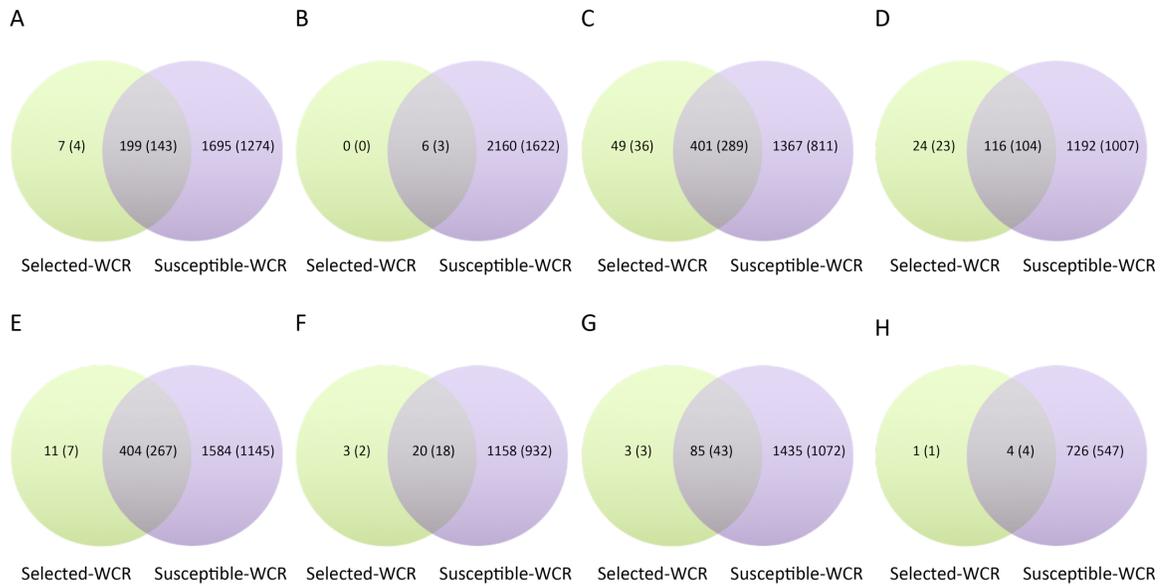
**Figure 2-1.** Contig length distribution of WCR (A) larval and (B) midgut transcriptome assemblies. The length of each contig has been converted to log scale.



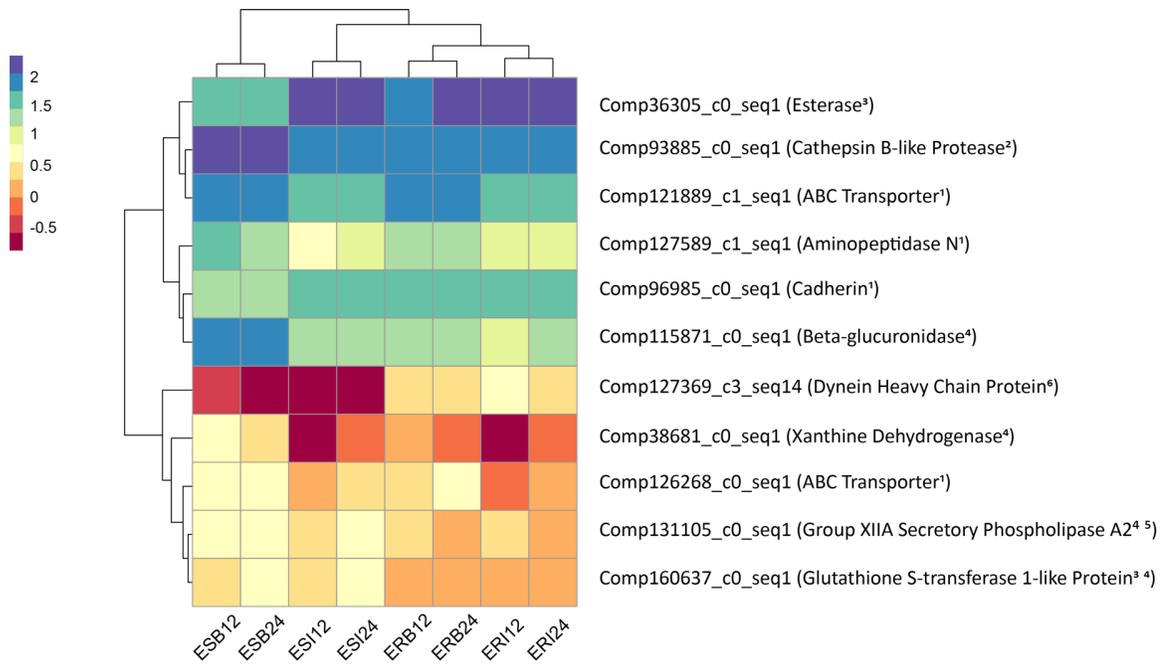
**Figure 2-2.** The species distribution of BLASTX top hits of WCR (A) larval and (B) midgut transcriptome assemblies.



**Figure 2-3.** Transcriptional response of neonate WCR allowed feed 12 or 24 hrs on eCry3.1Ab or non-Bt isoline maize seedlings compared to resistant neonates feeding on isoline seedlings. R: eCry3.1Ab resistant WCR; S: susceptible WCR; Bt: eCry3.1Ab transgenic seedlings; Isoline: non-Bt seedlings; Differential expression pattern: (A) larval response at 12 hours. (B) Larval response at 24 hours. (C) Midgut response at 12 hours. (D) Midgut response at 24 hours (adjust p-value <0.05, FDR test).



**Figure 2-4.** Differential expression pattern of resistant and susceptible WCR in whole larvae, or dissected midguts, when feeding on Cry3.1Ab transgenic maize seedlings vs. feeding on non-Bt isoline seedlings, for 12 or 24 hrs. In whole larvae (A, B, C, and D) at 12 hrs feeding up-regulated (A), and down-regulated (B) contigs, showing overlapping expression. In whole larvae at 24 hrs feeding up-regulated (C), and down-regulated (D). In dissected midgut tissues (E, F, G, and H) contigs up-regulated at 12 hr feeding (E), and down-regulated (F). At 24 hrs feeding contigs up-regulated (G) and down-regulated (H). Numbers in shown in parentheses are contigs with BLASTX annotation.



**Figure 2-5.** Expression levels of 11 potential eCry3.1Ab resistant related genes in 8 WCR midgut treatments (for treatment code see Table 2-1). Expression levels are quantified by count per million read (cpm). The candidate genes are categorized into: 1: potential Bt receptors; 2: digestive proteases; 3: detoxification enzymes; 4: enzymes involve in drug metabolism pathway; 5: enzymes involve in membrane related pathways; 6: other candidates.

## Bibliography

- Ball, H. J. and Weekman, G. T. (1962) Insecticide resistance in the adult western corn rootworm in Nebraska. *Journal of Economic Entomology* **55**: 439-441.
- Bonfini, A., Liu, X. and Buchon, N. (2016) From pathogens to microbiota: how *Drosophila* intestinal stem cells react to gut microbes. *Developmental & Comparative Immunology* **64**: 22-38.
- Bravo, A., Gill, S. S. and Soberón, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49**: 423-435.
- Bravo, A., Gómez, I., Conde, J., Muñoz-Garay, C., Sánchez, J., Miranda, R., *et al.* (2004) Oligomerization triggers binding of a *Bacillus thuringiensis* Cry1Ab pore-forming toxin to aminopeptidase N receptor leading to insertion into membrane microdomains. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1667**: 38-46.
- Caccia, S., Di Lelio, I., La Storia, A., Marinelli, A., Varricchio, P., Franzetti, E., *et al.* (2016) Midgut microbiota and host immunocompetence underlie *Bacillus thuringiensis* killing mechanism. *Proceedings of the National Academy of Sciences* **113**: 9486-9491.
- Castagnola, A. and Jurat-Fuentes, J. L. (2016) Intestinal regeneration as an insect resistance mechanism to entomopathogenic bacteria. *Current Opinion in Insect Science* **15**: 104-110.
- Chu, C.-C., Zavala, J. A., Spencer, J. L., Curzi, M. J., Fields, C. J., Drnevich, J., *et al.* (2015) Patterns of differential gene expression in adult rotation-resistant and wild-type western corn rootworm digestive tracts. *Evolutionary Applications* **8**: 692-704.
- Clark, T. L., Frank, D. L., French, B. W., Meinke, L. J., Moellenbeck, D., Vaughn, T. T., *et al.* (2012) Mortality impact of MON863 transgenic maize roots on western corn rootworm larvae in the field. *Journal of Applied Entomology* **136**: 721-729.
- Coates, B. S. (2014) Assembly and annotation of full mitochondrial genomes for the corn rootworm species, *Diabrotica virgifera virgifera* and *Diabrotica barberi* (Insecta: Coleoptera: Chrysomelidae), using next generation sequence data. *Gene* **542**: 190-197.
- Coates, B. S., Alves, A. P., Wang, H., Zhou, X., Nowatzki, T. M., Chen, H., *et al.* (2016) Quantitative trait locus mapping and functional genomics of an organophosphate resistance trait in the western corn rootworm, *Diabrotica virgifera virgifera*. *Insect Molecular Biology* **25**: 1-15.
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676.

- Eyun, S.-i., Wang, H., Pauchet, Y., ffrench-Constant, R. H., Benson, A. K., Valencia-Jiménez, A., *et al.* (2014) Molecular evolution of glycoside hydrolase genes in the western corn rootworm (*Diabrotica virgifera virgifera*). *PLoS ONE* **9**: e94052.
- ffrench-Constant, R. H., Daborn, P. J. and Goff, G. L. (2004) The genetics and genomics of insecticide resistance. *Trends in Genetics* **20**: 163-170.
- Flagel, L. E., Swarup, S., Chen, M., Bauer, C., Wanjugi, H., Carroll, M., *et al.* (2015) Genetic markers for western corn rootworm resistance to Bt toxin. *G3: Genes|Genomes|Genetics* **5**: 399-405.
- Frank, D. L., Zukoff, A., Barry, J., Higdon, M. L. and Hibbard, B. E. (2013) Development of resistance to eCry3.1Ab-expressing transgenic maize in a laboratory-selected population of western corn rootworm (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **106**: 2506-2513.
- Gahan, L. J., Gould, F. and Heckel, D. G. (2001) Identification of a gene associated with Bt resistance in *Heliothis virescens*. *Science* **293**: 857.
- Gahan, L. J., Pauchet, Y., Vogel, H. and Heckel, D. G. (2010) An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLOS Genetics* **6**: e1001248.
- Gassmann, A. J., Petzold-Maxwell, J. L., Keweshan, R. S. and Dunbar, M. W. (2011) Field-evolved resistance to Bt maize by western corn rootworm. *PLoS One* **6**: e22629.
- Gould, F. (1998) Sustainability of transgenic insecticidal cultivars: integrating pest genetics and ecology. *Annual Review of Entomology* **43**: 701-726.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644-652.
- Gunning, R. V., Dang, H. T., Kemp, F. C., Nicholson, I. C. and Moores, G. D. (2005) New resistance mechanism in *Helicoverpa armigera* threatens transgenic crops expressing *Bacillus thuringiensis* Cry1Ac toxin. *Applied and Environmental Microbiology* **71**: 2558-2563.
- Heckel, D. G., Gahan, L. J., Baxter, S. W., Zhao, J.-Z., Shelton, A. M., Gould, F., *et al.* (2007) The diversity of Bt resistance genes in species of Lepidoptera. *Journal of Invertebrate Pathology* **95**: 192-197.
- Hibbard, B. E., Clark, T. L., Ellersieck, M. R., Meihls, L. N., Khishen, A. A. E., Kaster, V., *et al.* (2010a) Mortality of western corn rootworm larvae on MIR604 transgenic maize roots: field survivorship has no significant impact on survivorship of F1 progeny on MIR604. *Journal of Economic Entomology* **103**: 2187-2196.

- Hibbard, B. E., Meihls, L. N., Ellersieck, M. R. and Onstad, D. W. (2010b) Density-dependent and density-independent mortality of the western corn rootworm: impact on dose calculations of rootworm-resistant Bt corn. *Journal of Economic Entomology* **103**: 77-84.
- Ihaka, R. and Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**: 299-314.
- Kanehisa, M. and Susumu, G. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**.
- Knowles, B. H. (1994) Mechanism of action of *Bacillus thuringiensis* insecticidal  $\delta$ -endotoxins. *Advances in Insect Physiology* **24**: 275-308.
- Langmead, B. and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357-359.
- Lawrence, M., Huber, W., PagÈs, H., Aboyoun, P., Carlson, M., Gentleman, R., *et al.* (2013) Software for computing and annotating genomic ranges. *PLoS Computational Biology* **9**: e1003118.
- Lee, J.-H., Shin, K.-S., Suh, S.-C., Rhim, S.-L., Lee, Y.-H., Lim, M.-H., *et al.* (2013) CryIIIa toxin gene expression in transgenic rice confers resistance to rice water weevil. *Plant Cell, Tissue and Organ Culture (PCTOC)* **115**: 243-252.
- Lefko, S. A., Nowatzki, T. M., Thompson, S. D., Binning, R. R., Pascual, M. A., Peters, M. L., *et al.* (2008) Characterizing laboratory colonies of western corn rootworm (Coleoptera: Chrysomelidae) selected for survival on maize containing event DAS-59122-7. *Journal of Applied Entomology* **132**: 189-204.
- Levine, E., Spencer, J. L., Isard, S. A., Onstad, D. W. and Gray, M. E. (2002) Adaptation of the western corn rootworm to crop rotation: evolution of a new strain in response to a management practice. *American Entomologist* **48**: 94-107.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Loseva, O., Ibrahim, M., Candas, M., Koller, C. N., Bauer, L. S. and Bulla Jr, L. A. (2002) Changes in protease activity and Cry3Aa toxin binding in the Colorado potato beetle: implications for insect resistance to *Bacillus thuringiensis* toxins. *Insect Biochemistry and Molecular Biology* **32**: 567-577.
- Ludwick, D. C. and Hibbard, B. E. (2016) Rootworm management: status of GM traits, insecticides and potential new tools. *CAB Reviews* **11**: 1-10.
- Ludwick, D. C., Meihls, L. N., Ostlie, K. R., Potter, B. D., French, L. and Hibbard, B. E. (2017) Minnesota field population of western corn rootworm (Coleoptera: Chrysomelidae)

shows incomplete resistance to Cry34Ab1/Cry35Ab1 and Cry3Bb1. *Journal of Applied Entomology* **141**: 28-40.

McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**: 4288-4297.

Meihls, L. N., Higdon, M. L., Ellersieck, M. and Hibbard, B. E. (2011) Selection for resistance to mCry3A-expressing transgenic corn in western corn rootworm. *Journal of Economic Entomology* **104**: 1045-1054.

Meihls, L. N., Higdon, M. L., Siegfried, B. D., Miller, N. J., Sappington, T. W., Ellersieck, M. R., *et al.* (2008) Increased survival of western corn rootworm on transgenic corn within three generations of on-plant greenhouse selection. *Proceedings of the National Academy of Sciences* **105**: 19177-19182.

Meinke, L. J., Siegfried, B. D., Wright, R. J. and Chandler, L. D. (1998) Adult susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to selected insecticides. *Journal of Economic Entomology* **91**: 594-600.

Miller, N. J. and Zhao, Z. (2015) Transcriptional responses of *Diabrotica virgifera virgifera* larvae to benzoxazinoids. *Journal of Applied Entomology* **139**: 416-423.

Pereira, A. E., Wang, H., Zukoff, S. N., Meinke, L. J., French, B. W. and Siegfried, B. D. (2015) Evidence of field-evolved resistance to bifenthrin in western corn rootworm (*Diabrotica virgifera virgifera* LeConte) populations in western Nebraska and Kansas. *PLOS ONE* **10**: e0142299.

Perlak, F. J., Stone, T. B., Muskopf, Y. M., Petersen, L. J., Parker, G. B., McPherson, S. A., *et al.* (1993) Genetically improved potatoes: protection from damage by Colorado potato beetles. *Plant Molecular Biology* **22**: 313-321.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., *et al.* (2005) InterProScan: protein domains identifier. *Nucleic Acids Research* **33**: W116-W120.

Rault, L. C., Siegfried, B. D., Gassmann, A. J., Wang, H., Brewer, G. J. and Miller, N. J. (2018) Investigation of Cry3Bb1 resistance and intoxication in western corn rootworm by RNA sequencing. *Journal of Applied Entomology* **142**: 921-936.

Rausell, C., Ochoa-Campuzano, C., Martínez-Ramírez, A. C., Bravo, A. and Real, M. D. (2007) A membrane associated metalloprotease cleaves Cry3Aa *Bacillus thuringiensis* toxin reducing pore formation in Colorado potato beetle brush border membrane vesicles. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1768**: 2293-2299.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140.

- Robinson, M. D. and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**: R25.
- Sayed, A., Nekl, E. R., Siqueira, H. A. A., Wang, H. C., Ffrench-Constant, R. H., Bagley, M., *et al.* (2007) A novel cadherin-like gene from western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae), larval midgut tissue. *Insect Molecular Biology* **16**: 591-600.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. and Zdobnov, E. M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210-3212.
- Storer, N. P., Babcock, J. M. and Edwards, J. M. (2006) Field measures of western corn rootworm (Coleoptera: Chrysomelidae) mortality caused by Cry34/35Ab1 proteins expressed in maize event 59122 and implications for trait durability. *Journal of Economic Entomology* **99**: 1381-1387.
- Tabashnik, B. E. and Gould, F. (2012) Delaying corn rootworm resistance to Bt corn. *Journal of Economic Entomology* **105**: 767-776.
- Walters, F. S., deFontes, C. M., Hart, H., Warren, G. W. and Chen, J. S. (2010) Lepidopteran-active variable-region sequence imparts coleopteran activity in eCry3.1Ab, an engineered *Bacillus thuringiensis* hybrid insecticidal protein. *Applied and Environmental Microbiology* **76**: 3082-3088.
- Wang, H., Eyun, S.-i., Arora, K., Tan, S., Gandra, P., Moriyama, E., *et al.* (2017) Patterns of gene expression in western corn rootworm (*Diabrotica virgifera virgifera*) neonates, challenged with Cry34Ab1, Cry35Ab1 and Cry34/35Ab1, based on next-generation sequencing. *Toxins* **9**: 124.
- Wangila, D. S., Gassmann, A. J., Petzold-Maxwell, J. L., French, B. W. and Meinke, L. J. (2015) Susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to Bt corn events. *Journal of Economic Entomology* **108**: 742-751.
- Wei, J., Liang, G., Wang, B., Zhong, F., Chen, L., Khaing, M. M., *et al.* (2016) Activation of Bt protoxin Cry1Ac in resistant and susceptible cotton bollworm. *PLoS ONE* **11**: e0156560.
- Zhao, J.-Z., Oneal, M. A., Richtman, N. M., Thompson, S. D., Cowart, M. C., Nelson, M. E., *et al.* (2016) mCry3A-Selected western corn rootworm (Coleoptera: Chrysomelidae) colony exhibits high resistance and has reduced binding of mCry3A to midgut tissue. *Journal of Economic Entomology* **109**: 1369-1377.
- Zhou, X., Lindsay, H. and Robinson, M. D. (2014) Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Research* **42**: e91-e91.

Zukoff, S. N., Ostlie, K. R., Potter, B., Meihls, L. N., Zukoff, A. L., French, L., *et al.* (2016) Multiple assays indicate varying levels of cross resistance in Cry3Bb1-selected field populations of the western corn rootworm to mCry3A, eCry3.1Ab, and Cry34/35Ab1. *Journal of Economic Entomology* **109**: 1387-1398.

### **Chapter 3**

The detection of alternative splicing in *Diabrotica virgifera virgifera* LeConte in association with eCry3.1Ab resistance

## Abstract

Alternative splicing is one of the major mechanisms that increase transcriptome diversity in eukaryotes. The phenomenon is observed in many insect species that have gained resistance to pesticides and Bt toxins. In western corn rootworm (*Diabrotica virgifera virgifera* LeConte), neither alternative splicing nor its role in resistance to Bt has been studied. Here we present the transcriptome-wide alternative splicing patterns of western corn rootworm midgut using a comprehensive approach that combines RNA-seq and third-generation single molecule isoform sequencing. We found that 63.11% of multi-exon genes were alternatively spliced, consistent with the high transposable element content in the rootworm genome. One of the alternative splicing events we identified was a novel peritrophic matrix protein with two alternative splicing isoforms. Analysis of differential exon usage showed that in eCry3.1Ab-resistant western corn rootworm, expression of one isoform is significantly higher than in the susceptible colony, while no significant differences between colonies were observed with the other isoform. Our results therefore provide the first survey of alternative splicing in western corn rootworm and suggested that the observed alternatively spliced isoforms of peritrophic matrix protein might be associated with eCry3.1Ab resistance.

## Introduction

The western corn rootworm (WCR), *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae) is one of the most destructive maize pests in the US Corn Belt. WCR demonstrates genetic plasticity and has developed resistance to many control strategies, including transgenic maize expressing rootworm-targeting Bt toxins Cry3Bb1, mCry3A, eCry3.1Ab, and Cry34/35Ab1. Resistance to Bt maize has been selected in laboratory colonies (Dietloff et al. 2016; Frank et al. 2013; Meihls et al. 2011; Meihls et al. 2008). In the field, failure of Bt-maize lines to control rootworm has been reported in multiple states (Gassmann et al. 2011; Ludwick et al. 2017; Wangila et al. 2015; Zuckoff et al. 2016). This has highlighted the importance of managing and understanding the mechanisms of Bt resistance in WCR.

Mechanisms of Bt resistance in WCR have not been well elucidated at the gene level. High-throughput RNA sequencing (RNA-seq) provides a powerful tool to analyze gene expression profiles, especially to identify differentially expressed genes that are associated with Bt resistance (Rault et al. 2018; Wang et al. 2017; Zhao et al. 2019). The transcript sequences generated in these studies were either from expressed sequence tags (EST) or assembled *de novo* from short-length RNA-seq reads. In other studies, genetic markers were developed for analyzing quantitative traits loci (QTL) that linked to Bt resistance in WCR (Flagel et al. 2015). However, gene expression regulation and genomic variation may not be the only causes of resistance. Post-transcriptional processes such as alternative splicing (AS) may create isoforms that confer additional functions, i.e. Bt resistance, in WCR.

AS is a common phenomenon in eukaryotic organisms. A multi-exon gene can use different splice sites to create multiple transcripts with different exon combinations. In human, about 95% of multi-exon genes produce more than one transcript by AS (Pan et al. 2008). In insects, AS is involved in many physiological processes, including sex determination (Salz 2011), social behaviors (Jarosch et al. 2011), molting (Abdel-Banat and Koga 2002) and immune responses (Riddell et al. 2014). Resistance to insecticides can be achieved by AS of target receptors, e.g. sodium channels, which produce functional receptors with diverse structures. Sodium channel isoforms can vary in sensitivity to insecticides (Dong 2007). The glutathione transferases (GSTs), a class of detoxification enzyme involved in resistance to insecticides, are alternatively spliced in some mosquito species (Kasai et al. 2009; Strode et al. 2008). Recent studies on the pink bollworm (*Pectinophora gossypiella*) demonstrated that splice variants of a cadherin receptor *PgCad1* and an ATP-binding cassette transporter (ABC transporter) *PgABCA2* cause resistance to Cry1Ac and Cry2Ab (Fabrick et al. 2014; Mathew et al. 2018). Similarly, AS of corn earworm (*Helicoverpa armigera*) ABC transporter *HaABCC2* was shown to associate with Cry1Ac resistance (Xiao et al. 2014). Both revealed that AS is an evolutionary strategy and potential source for insects to adapt to environmental stresses, including Bt intoxication.

Few studies have focused on AS in WCR. Part of the reason was the lack of an annotated genome and other tools that would enable detection of genome-wide AS. The *de novo* assembled transcripts from RNA-seq do not provide an accurate representation of isoforms. In recent years, single molecule real-time (SMRT) PacBio sequencing technology has been introduced (Pacific Biosciences, Menlo Park, CA) (Rhoads and Au

2015). This technology generates high-quality, ultra-long sequencing reads, sometimes greater than 10 kb. Isoform sequencing (Iso-seq) is the application of PacBio sequencing to transcriptomes, which can produce full-length transcripts and splicing variants without assembly artifacts such as chimeras (Gonzalez-Garay 2016). Iso-seq provides simplified and reliable approaches for detecting AS and improves the current genome annotation of non-model organisms like WCR.

We report the first genome-wide survey of splice isoforms in WCR using PacBio Iso-seq. The isoform profile was generated from a lab-selected eCry3.1Ab-resistant colony and its non-selected susceptible counterpart colony (Frank et al. 2013). To detect the AS events associated with Bt resistance, we conducted a comprehensive analysis from a previous Illumina-based RNA-seq study of midgut response to eCry3.1Ab, with midgut transcripts sequenced by PacBio Iso-seq. We were able to identify differentially spliced transcripts between eCry3.1Ab-resistant and susceptible WCR colonies, with or without intoxication from the eCry3.1Ab. Our results revealed that colony-specific splicing events are specific to Bt intoxication and exposure time. Among these we identified and characterized a peritrophic matrix protein (PMP) with a colony-specific splicing pattern, which may be associated with eCry3.1Ab resistance in WCR larvae.

## **Material and methods**

### **Insects and data source**

The non-diapause eCry3.1Ab-resistant WCR colony (NB) and susceptible colony (NI) are from continuous lab-selection experiments (Frank et al. 2013). Another non-diapause eCry3.1Ab-susceptible WCR colony (SD) was obtained from USDA-ARS at Brookings, SD. The RNA-seq experiments were conducted on NB and NI. The insect

feeding assay, midgut dissection, RNA-extraction and sequencing, and data trimming have been previously described in Chapter 2 (Zhao et al. 2019). The *D. v. virgifera* reference genome assembly (version 1), which was sequenced and assembled from WCR strain PED-12, was downloaded from NCBI genome database ([https://www.ncbi.nlm.nih.gov/assembly/GCA\\_003013835.1/](https://www.ncbi.nlm.nih.gov/assembly/GCA_003013835.1/)).

### **Bioassay, Iso-seq library preparation, and sequencing**

Prior to creating libraries for Iso-seq, the same insect feeding assay that was applied in the RNA-Seq experiment was carried out on NB and NI colonies (Table 3-1). Total RNA was extracted from 30-40 dissected neonate midguts using the Direct-zol RNA Mini Prep kit (Zymo Research, Irving, CA) with DNase treatment (Thermo Fisher, Waltham, MA). To increase the sequence coverage of Iso-seq we pooled an equal amount of total RNA from each sample. Iso-seq library preparation and PacBio sequencing were carried out by Genewiz LLC (South Plainfield, NJ). The library was loaded and sequenced on three cells of PacBio Sequel sequencer to increase the sequence depth. Size-selection was not implemented, according to factory recommendation.

### **Iso-seq data processing and transcript analysis**

The PacBio SMRTLink (v 5.0.1) Iso-seq pipeline was employed to extract sequences from sequencer outputs. The pipeline contains three steps. The first circular consensus sequence (CCS) step generates CCS reads from sequencer outputs. The second classification step identifies full-length transcripts. The third clustering step further clusters and polishes full-length sequences using ICE and Quiver algorithms, and partitions the full-length sequences into high-quality and low-quality categories. SMRTLink analysis was followed by post-analysis pipelines that further collapsed

sequences into genes and isoforms. To maximize the sensitivity of detecting gene and isoforms, we merged both high-quality and low-quality full-length transcripts. The post-analysis was initiated by aligning full-length transcripts to the WCR genome assembly using GMAP (version 2014-08-19) (Wu and Watanabe 2005). The aligned transcripts were further processed using Cupcake ToFU ([https://github.com/Magdoll/cDNA\\_Cupcake/wiki/Cupcake-ToFU:-supporting-scripts-for-Iso-Seq-after-clustering-step](https://github.com/Magdoll/cDNA_Cupcake/wiki/Cupcake-ToFU:-supporting-scripts-for-Iso-Seq-after-clustering-step)). For transcripts that failed to align to the genome, we first removed the duplicated sequences using a custom Perl script. Then a reference-free collapsing approach was accomplished using Cogent (<https://github.com/Magdoll/Cogent/wiki/Running-Cogent>). The unaligned transcripts were classified as singlets or clustered into gene families. Each gene family was further collapsed into one or several full-length unique transcript models, which predicted the isoform pattern of each gene.

To evaluate the completeness and coverage of Iso-seq we applied BUSCO (version 3) (Simão et al. 2015) with the lineage “insect\_obd9” separately on (1) isoforms of genes that aligned to genome assembly, (2) collapsed unaligned transcripts that include gene models and singlets, and (3) the merger of above two sets.

### **Identification of colony-specific spliced genes**

We first aligned all the RNA-seq reads to the WCR reference genome using TopHat (version 2.0.12) (Trapnell et al. 2009). Due to the lack of annotation of genomic features, we created an annotation file based on Iso-seq alignments as follows. We used a custom Perl script to convert the Cupcake ToFU output of gene structure of aligned transcripts (in GFF format) to a genomic feature file in Ensembl GTF format, and then

we assigned gene and transcript with unique IDs with arbitrary prefixes. Based on RNA-seq we compared the differences in alternative splicing patterns between NB and NI colonies. The DEXSeq package (Anders et al. 2012) was then employed to test changes in exon usage of each gene between NB and NI samples and report genes with significantly different exon usage (DEU). The DEU represents the relative abundance of each exon within a gene, which demonstrates the differences of isoform pattern between colonies.

### **Functional annotation of Iso-seq transcriptome and GO enrichment of colony-specific AS genes**

Due to the lack of functional annotation of the WCR genome, we annotated the Iso-seq transcriptome using the Blast2GO (version 5.2.5) (Conesa et al. 2005) standard annotation pipeline. The isoforms were first searched against the NCBI non-redundant protein database (nr) using BLASTX. The gene ontology (GO) terms were obtained by mapping BLASTX results to the Gene Ontology annotation database (GOA) (Huntley et al. 2014). Then we applied Blast2GO default annotation rules to the above GO terms, resulting in the first GO annotation dataset. We estimated the domain/motif composition of isoforms using InterProScan. The GO terms reported from InterProScan were merged with existing GO annotations to create the final GO annotation dataset.

To better understand the function of colony-specific spliced genes we applied GO enrichment analysis using the R package topGO (Alexa and Rahnenfuhrer 2010). We tested if any GO terms from putative colony-specific AS genes were over-represented compared to the total annotated gene set from WCR genome from Iso-seq. GO annotations of each gene were obtained by merging the annotated GO terms from its

isoforms. Fisher's exact tests were applied separately on three categories: biological process (BP), molecular function (MF), and cellular component (CC).

### **Characterization of a colony-specific spliced novel peritrophic matrix protein**

The mutually exclusive exon usage pattern of novel peritrophic membrane protein (PMP, gene alias PB8671) was tested by PCR using isoform-specific primers (Table 3-1). Quantitative RT-PCR (qRT-PCR) was performed on NB and SD colonies after the same bioassay, midgut dissection, and RNA extraction described above. Three independent studies were conducted using three cohorts of the above colonies as biological replications. The cDNA is synthesized from midgut RNA using Omniscript RT kit and followed by real-time PCR using QuantiFast SYBR Green PCR Kit (Qiagen, Germantown, MD). Two sets of primers were designed to quantify the expression of PB8671 in total or the specific isoform PB8671.2 alone. A housekeeping gene elongation factor-1 $\alpha$  (EF-1 $\alpha$ ) was incorporated as an internal control (Table 3-2) (Rodrigues et al. 2014). The relative expression levels were determined by the  $2^{-\Delta\Delta CT}$  method (Livak and Schmittgen 2001). ANOVA tests were implemented to  $\Delta Ct$  (Ct of target gene minus Ct of EF-1 $\alpha$ ) of each sample to compare expression among treatment.

## **Results**

### **Overview of WCR transcriptome and AS pattern from PacBio Iso-seq**

We obtained 22,489,178 reads with good quality (subreads) from three Iso-seq Sequel sequencer cells. From the SMRTLink Iso-seq analysis pipeline, the CCS step claimed 1,046,323 CCS reads. The classification step identified 745,428 full-length non-chimeric sequences. The clustering step categorized 53,394 high-quality full-length sequences and 322,095 low-quality full-length sequences. A total of 36,184 high quality

(67.77%) and 188,780 low quality (58.61%) full-length transcripts could be aligned to the WCR reference genome.

From transcripts that were aligned to the WCR reference genome, the post-analysis software Cupcake ToFU identified 12,023 gene loci, from which 32,410 total transcripts were produced. Among the gene loci were 3,848 single-exon genes. For 8,175 multi-exon genes, 63.11% (5195) produced more than one transcript. From a total of 150,525 unaligned transcripts, 9,900 duplicated sequences were removed. The software Cogent generated a total number of 4,963 gene families with 24,325 transcripts models. There were 42,699 transcripts that did not belong to or form any gene family. These were claimed as singlets.

From BUSCO reports, the set of isoforms from gene loci contained only 60.6% complete conserved single-copy orthologs, while the unaligned set of gene models and singlets contained 50.7%. The merger of the above two sets resulted in an increased completeness to 79.9%. This result demonstrated that many coding genes are still not incorporated in the current reference genome assembly cited above. Transcriptome coverage was improved when unaligned transcripts were included. Compared with *de novo* transcriptome assembly from RNA-seq (Zhao et al. 2019), the Iso-seq transcriptome had a lower coverage, but an improved continuity measured as the minimum sequence length that covered 50% of the assembly (N50) (Table 3-4).

### **Functional annotation of WCR transcriptome**

Blast2GO functional annotation of Iso-seq transcripts showed a total of 78.0% and 60.1% of aligned and unaligned transcripts with significant BLASTX hits (E-value < 1e-3). Of these, 18,185 aligned (56.1%) and 20,627 unaligned (30.8%) transcripts were

annotated with GO terms. The level-2 GO annotation distribution showed a very similar pattern between aligned and unaligned transcripts (Figure 3-1). The most abundant molecular function and cellular component GO terms for both transcripts were “ATP-binding” (GO:0005524) and “integral component of membrane” (GO:0016021). In the biological process category, “oxidation-reduction process” (GO:0055114) and “translation” (GO:0006412) were the most abundant GO terms for aligned and unaligned transcripts, respectively. The results concurred with the physiological function of insect midgut in digestion and nutrient uptake.

With the functional annotation we were able to identify the splicing pattern of potential Bt receptors, i.e. cadherin and aminopeptidase N. The cadherin (EF531715), previously identified in WCR by Sayed et al. (2007), had 17 different AS isoforms (PB112) in our results. We found another putative DE-cadherin (PB4589) gene with 3 isoforms, and 2 separate putative protocadherin genes (PB1648 and PB127) with one isoform each from WCR genome. We also identified four different aminopeptidase N genes, each with three to five isoforms (Table 3-4).

### **Colony-specific alternative splicing of WCR**

Using RNA-seq data from NB and NI, we compared the AS patterns between NB and NI colonies under different experimental conditions. While setting the adjusted p-value cutoff at 0.1, as suggested by DEXSeq tutorial (Reyes et al. 2012), DEXSeq detected 39 and 52 genes with significant DEU between colonies when feeding on eCry3.1- and isoline maize respectively for 12 hours. After feeding for 24 hours, there were 55 and 105 genes were detected with significant DEU between colonies. Only five genes had DEU between NB and NI regardless of diet type and time (Figure 3-2),

however, two of which had nearly identical DEU patterns between colonies (Table 3-5). To understand the biological function associated with colony-specific spliced genes we applied GO enrichment analysis using Fisher's exact test. The enriched GO terms were also very specific to food type and time frame. None of the GO terms of any category were present as significantly enriched for both isoline and Bt-maize feeding at 2 time points (Figure 3-3).

When both resistant and susceptible larvae were fed with Bt-maize, the most significantly enriched GO terms in the molecular function (MF) category of differentially spliced were 'signaling receptor activity' (GO:0038023) and 'signal transducer activity' (GO:0004871) at 12 hours (Table 3-6A), indicating that colony-specific splicing occurred in genes involved in signaling pathways. At 24 hours the top MF GO terms were 'peptidase activity, acting on L-amino acid peptides' (GO:0070011) and 'hydrolase activity' (GO:0016787) (Table 3-6D). Enriched GO terms in the cellular component (CC) category also revealed the trend that colony-specific spliced genes were switching over time from the extracellular region to intracellular and organelle regions (Table 3-6).

### **Characterization of the colony-specific spliced novel peritrophic membrane protein**

The colony-specific splicing patterns are likely due to selection pressure imposed by eCry3.1Ab. One gene with consistent colony-specific splicing pattern mentioned above was identified as a novel peritrophic membrane protein (PMP, gene ID in GTF: PBDG00000008671, gene alias in Iso-seq: PB8671). The Iso-seq identified a gene with two isoforms that had mutually exclusive exon usage, with transcript PB8671.1 expressing exon E003 and transcript PB8671.2 expressing exons E004 with E005 (Figure 3-4A and Table 3-7). Detailed analysis of the genomic locus of PB8671 showed that this

gene is on the negative strand of genome region shown in Figures 3-4A and 3-7. Transcription can be initiated from either exon E005 or E003. PCR amplification of the regions covering exon boundary further confirmed that E003 is skipped in isoform PB8671.2 (Figure 3-4B). Sequence analysis predicted that both isoforms had complete open reading frames (ORFs). The sizes of the corresponding peptides were either 99 amino acids (AA) (PB8671.1) or 112 AA (PB8671.2), each with one conserved 6-cysteine ChtBD2-type chitin binding domain (pfam01607) (Figure 3-5). A 16 AA N-terminal signal peptide is predicted from both isoforms using SignalP-5.0 (Almagro Armenteros et al. 2019) (Figure 3-6). The DEXSeq analysis demonstrated exon usage of E004 and E005 were significantly increased and the usage of E003 was slightly decreased in the NB colony, compared to NI (Figure 3-7). That suggested the transcription of isoform PB8671.2 was increased at both 12 and 24 hours in eCry3.1Ab1-resistant WCR, regardless of the presences of toxin.

The expression of isoform PB8671.2 and PB8671 in total were further analyzed by qRT-PCR from the NB and SD colonies. The relative transcript expression levels of PB8671 were slightly higher in NB cohorts than SD. The difference was significant at 12 hours when they were feeding on eCry3.1Ab maize (Figure 6B & 6D). In contrast, the relative expression level difference of isoform PB8671.2 was insignificant and less consistent within and between colonies. The NB-1 cohort expressed more PB8671.2 than any of SD cohorts in any condition. The NB-2 expressed more PB8671.2 than SD cohorts with both feeding types at 12 hours and isoline feeding at 24 hours. The NB-3 only showed increased PB8671.2 expression after feeding on eCry3.1Ab maize for 24 hours (Figure 6A & 6C). We also found expressions of either PB8671.2 or PB8671 in total

were reduced on eCry3.1Ab maize feeding than isoline feeding WCR, but the differences were insignificant.

## Discussion

The mode of action of Bt toxin and the mechanism of Bt resistance has been extensively studied in lepidopteran insects (Bravo et al. 2007; Bravo and Soberón 2008; Heckel et al. 2007). In WCR, genomics research is restricted by a complicated and incomplete reference genome assembly. The common strategy to overcome this limitation is quantification of genes on a *de novo* assembled transcriptome. The third-generation PacBio sequencing technology generates ultra-long reads, which is a powerful tool to improve genome sequencing, genome annotation, novel gene discovery, and isoform presentation for non-model organisms like WCR. Thus, we first implemented an isoform survey on WCR using PacBio Iso-seq and did a comprehensive analysis in combination with RNA-seq, particularly the splicing profiles in association with eCry3.1Ab resistance. The goal of performing Iso-seq was to provide an accurate pattern of gene structure and genomic features of WCR, while RNA-seq data had a higher sequence depth, which was able to quantify the usage of particular exons in specific samples. The genome sequence of WCR allowed us to improve the quality of gene construction and quantification of exon usage. However, a large number of contigs failed to align to the genome, indicating large gaps exist in the current genome assembly.

In this study, we carried out differential exon usage analysis on gene loci identified in the reference genome. We discovered about 63.11% of multi-exon genes in WCR had an AS event, which was significantly greater than for *Drosophila* EST data (19%) (Kim et al. 2006). The result is consistent with a large WCR genome with many

transposable elements (TEs) (Coates et al. 2012), a possible mechanism for the emergence of AS. When comparing AS profiles by exon usage between eCry3.1Ab-resistant and susceptible WCR, the colony-specific isoforms varied among treatments and feeding times. Only a few genes showed persistent colony-specific splicing pattern upon Bt intoxication (eCry3.1Ab or isoline maize). GO enrichment results supported the same pattern. Compared to differentially expressed genes obtained from the RNA-seq dataset described in Chapter 2, genes with colony-specific splicing patterns were highly specific to experimental conditions. The results indicate that AS is served as the fine tune mechanism in addition to gene expression.

Mis-spliced isoforms of Bt receptors, i.e. cadherin or ABC transporters, have been found in several Lepidopteran resistance species (Fabrick et al. 2014; Mathew et al. 2018; Xiao et al. 2014). However, the splicing patterns in WCR have not been studied. Using Iso-seq we identified that some putative Bt receptors, i.e. cadherin (EF531715) (Sayed et al. 2007) and aminopeptidase N, were alternatively spliced. However their patterns are not colony-specific.

In insects, PMP is a group of chitin-binding proteins expressed only in midgut. The common feature of PMP is one or more CBD (Tetreau et al. 2015). Although the peritrophic matrix is a barrier to pathogens, Bt toxins are able to pass through it and reach to the epithelial cells. The proposed functions include binding and cross-linking chitin fibrils to form peritrophic membrane. Studies have shown that in Lepidoptera Cry toxins are able to bind PMP and this binding may contribute to Bt susceptibility in some insects (Rees et al. 2009; Valaitis and Podgwaite 2013). The binding of eCry3.1Ab to PMP in WCR should be studied. As a potential Bt receptor, some APNs are also located in the

peritrophic membrane (Campbell et al. 2008; Hu et al. 2012). Further studies are necessary to explore whether splicing of PMP affects peritrophic membrane stability and interaction with APN. In the presented study, we reported a novel PMP that consists of two AS isoforms. The PB8671.2 isoform has an increased expression in resistant WCR. The qRT-PCR showed up-regulation of PMP in total in resistance WCR. The consistent biased-splicing pattern and up-regulation in the resistant colony might be the consequence of selection on eCry3.1Ab maize. However, due to the short length of the alternatively spliced region, we lacked the primers that specifically target isoform PB8671.1, which caused difficulty in directly comparing PB8671.1 with PB8671.2. The specific primers of PB8671.2 were located at 5'-end with less optimal efficacy, which resulted in inconsistent quantification.

As conclusion, we presented a survey of alternative splicing patterns in WCR using a combination of RNA-seq and PacBio Iso-seq. The full-length assembly-free transcriptome sequences served as a source for annotating gene structures. We also identified genes with colony-biased splicing pattern in eCry3.1Ab resistant NB colony, and characterized a new PMP isoform from among them. The functions and roles in Bt resistance are yet to be studied.

## **Acknowledgements**

We thank Julie Barry for rearing WCR colonies and Justin Le Tourneau for software and server support. Deborah Triant and Adriano Pereira offered great suggestions for lab and data analysis.

**Table 3-1** The bioassay treatment of WCR neonate larvae before RNA-extraction and transcriptome PacBio Iso-seq.

<b>WCR Colony</b>	<b>Maize type</b>	<b>Feeding time (hours)</b>
NB	eCry3.1Ab-mazie	12
NB	Isoline	12
NB	eCry3.1Ab-mazie	24
NB	Isoline	24
NI	eCry3.1Ab-mazie	12
NI	Isoline	12
NI	eCry3.1Ab-mazie	24
NI	Isoline	24

**Table 3-2.** The PCR primers for validating alternative splicing pattern. Anticipated amplicon size for PB8671.2 is (A) 139 bp with skipped E003 or (B) 244 bp without exon skipping.

<b>Isoform</b>	<b>Location (exon)</b>	<b>Direction</b>	<b>Primer sequence</b>	<b>Anticipated amplicon length</b>
PB8671.1	E003	Forward	TGAAACCGCTTTTCAATACCCAG	143 bp
	E002	Reverse	TTCCAGGAGCACAGTGGAAT	
PB8671.2	E004	Forward	GCAGCACCAACTCATAGAC	139 (A) or 244 (B) bp
	E002	Reverse	GAGCACAGTGGAATTGGAAG	

**Table 3-3.** The primer sequences used for qRT-PCR. The housekeeping elongation factor 1 $\alpha$  (EF-1  $\alpha$ ) were used as internal reference.

Target	Primer name	Direction	Primer sequence
Elongation factor 1 $\alpha$ (housekeeping gene)	EF_1a_F	Forward	ACCAGATTTGATGGCTTTGG
	EF_1a_R	Reverse	CACCCAGAGGAGCTTCAGAC
PMP isoform PB8671.2	PMP_2_F	Forward	ACAAGTGGTTCTTCTCTGCG
	PMP_2_R	Reverse	TGAAGGTCTATGAGTTGGTGCT
PMP PB8671 total	PMP_T_F	Forward	GCACCTCCAGTCTACGAACC
	PMP_T_R	Reverse	ACGGTCTCCACAGTCAACATT

**Table 3-4.** Summary of BUSCO analysis of isoform from Iso-seq. The unaligned isoforms consist of singlets and transcript models of putative gene families of Cogent genome-free collapse. Total transcriptome is the merge of aligned and unaligned isoforms.

	<b>Aligned isoforms</b>	<b>Unaligned isoforms*</b>	<b>Total transcriptome</b>	<b><i>de novo</i> assembly from RNA-seq</b>
<b>Contig Number</b>	32,410	67,024	99,434	226,115
<b>Completeness</b>	60.60%	50.70%	79.90%	96.60%
<b>Single copy</b>	39.10%	32%	44.70%	65.40%
<b>Duplicated</b>	21.50%	18.70%	35.20%	31.20%
<b>Fragmented</b>	6.50%	7.30%	4.90%	1.70%
<b>Missing</b>	32.90%	42%	15.20%	1.70%
<b>N50</b>	2,072	2,479	2,358	1,351

**Table 3-5.** The number of isoforms of putative Bt receptor gene cadherin and aminopeptidase N.

<b>Iso-seq gene alias ID</b>	<b>Descripton</b>	<b>Number of isoforms</b>
PB112	Cadherin-like protein	17
PB4589	DE-cadherin-like protein	3
PB827	Protocadherin	1
PB1648	Protocadherin	1
PB4933	Aminopeptidase N	5
PB4990	Aminopeptidase N	3
PB7332	Aminopeptidase N	4
PB10470	Aminopeptidase N	7

**Table 3-6.** The top enriched GO terms from genes with colony-specific splicing patterns.

The ranks are based on p-value of classic Fisher test in R topGO package. (A) Molecular function (MF) terms of 12 hours; (B) Cellular component (CC) terms of 12 hours; (C) Biological process terms at 12 hours; (D) MF terms of 24 hours; (E) CC terms of 12 hours; (F) BP terms at 12 hours

**A.**

Rank	GO ID	GO Term	Number of Genes	Fisher p-value
1	GO:0038023	Signaling receptor activity	52	0.0057
2	GO:0004871	Signal transducer activity	56	0.007
3	GO:0004777	Succinate-semialdehyde dehydrogenase	1	0.0071
4	GO:0009013	Succinate-semialdehyde dehydrogenase	1	0.0071
5	GO:0050104	L-gulonate 3-dehydrogenase activity	1	0.0071
6	GO:0004872	Receptor activity	61	0.0088
7	GO:0060089	Molecular transducer activity	61	0.0088
8	GO:0008379	Thioredoxin peroxidase activity	2	0.0141
9	GO:0008474	Palmitoyl-(protein) hydrolase activity	2	0.0141
10	GO:0098599	Palmitoyl hydrolase activity	2	0.0141
11	GO:0003857	3-Hydroxyacyl-CoA dehydrogenase activity	3	0.0211
12	GO:0016019	Peptidoglycan receptor activity	3	0.0211
13	GO:0043141	ATP-dependent 5'-3' DNA helicase activity	3	0.0211
14	GO:0033961	cis-Stilbene-oxide hydrolase activity	4	0.028
15	GO:0043139	5'-3' DNA helicase activity	4	0.028

**B.**

Rank	GO ID	GO Term	Number of Genes	Fisher p-value
1	GO:0005576	Extracellular region	190	0.0034
2	GO:0005615	Extracellular space	82	0.0038
3	GO:0044421	Extracellular region part	98	0.0073
4	GO:0032580	Golgi cisterna membrane	2	0.016
5	GO:0097255	R2TP complex	2	0.016
6	GO:0000221	Vacuolar proton-transporting V-type ATPase	3	0.0239
7	GO:0005869	Dynactin complex	4	0.0317
8	GO:0000812	Swr1 complex	5	0.0395
9	GO:0005618	Cell wall	5	0.0395
10	GO:0005801	cis-Golgi network	5	0.0395
11	GO:0009274	Peptidoglycan-based cell wall	5	0.0395
12	GO:0005885	Arp2/3 protein complex	6	0.0472
13	GO:0009328	Phenylalanine-tRNA ligase complex	6	0.0472
14	GO:0030312	External encapsulating structure	6	0.0472

**C.**

Rank	GO ID	GO Term	Number of Genes	Fisher p-value
1	GO:0045321	Leukocyte activation	8	0.0017
2	GO:0001775	Cell activation	9	0.0021
3	GO:0035107	Appendage morphogenesis	11	0.0032
4	GO:0035114	Imaginal disc-derived appendage morphogenesis	11	0.0032
5	GO:0035120	Post-embryonic appendage morphogenesis	11	0.0032
6	GO:0048736	Appendage development	11	0.0032
7	GO:0048737	Imaginal disc-derived appendage development	11	0.0032
8	GO:0007560	Imaginal disc morphogenesis	12	0.0039
9	GO:0048563	Post-embryonic animal organ morphogenesis	12	0.0039
10	GO:0007552	Metamorphosis	15	0.0061
11	GO:0009886	Post-embryonic animal morphogenesis	15	0.0061
12	GO:0048569	Post-embryonic animal organ development	15	0.0061
13	GO:0048707	Instar larval or pupal morphogenesis	15	0.0061
14	GO:0060562	Epithelial tube morphogenesis	15	0.0061
15	GO:0050776	Regulation of immune response	16	0.0069

**D.**

Rank	GO ID	GO Term	Number of genes	Fisher p-value
1	GO:0070011	Peptidase activity, acting on L-amino acid peptides	277	0.0025
2	GO:0016787	Hydrolase activity	1251	0.0026
3	GO:0008233	Peptidase activity	301	0.0042
4	GO:0003726	Double-stranded RNA adenosine deaminase activity	1	0.0088
5	GO:0004818	Glutamate-tRNA ligase activity	1	0.0088
6	GO:0004819	Glutamine-tRNA ligase activity	1	0.0088
7	GO:0008073	Ornithine decarboxylase inhibitor activity	1	0.0088
8	GO:0008251	tRNA-specific adenosine deaminase activity	1	0.0088
9	GO:0035620	Ceramide transporter activity	1	0.0088
10	GO:0042979	Ornithine decarboxylase regulator activity	1	0.0088
11	GO:0046624	Sphingolipid transporter activity	1	0.0088
12	GO:0046625	Sphingolipid binding	1	0.0088
13	GO:0070273	Phosphatidylinositol-4-phosphate binding	1	0.0088
14	GO:0097001	Ceramide binding	1	0.0088
15	GO:1902387	Ceramide 1-phosphate binding	1	0.0088

**E.**

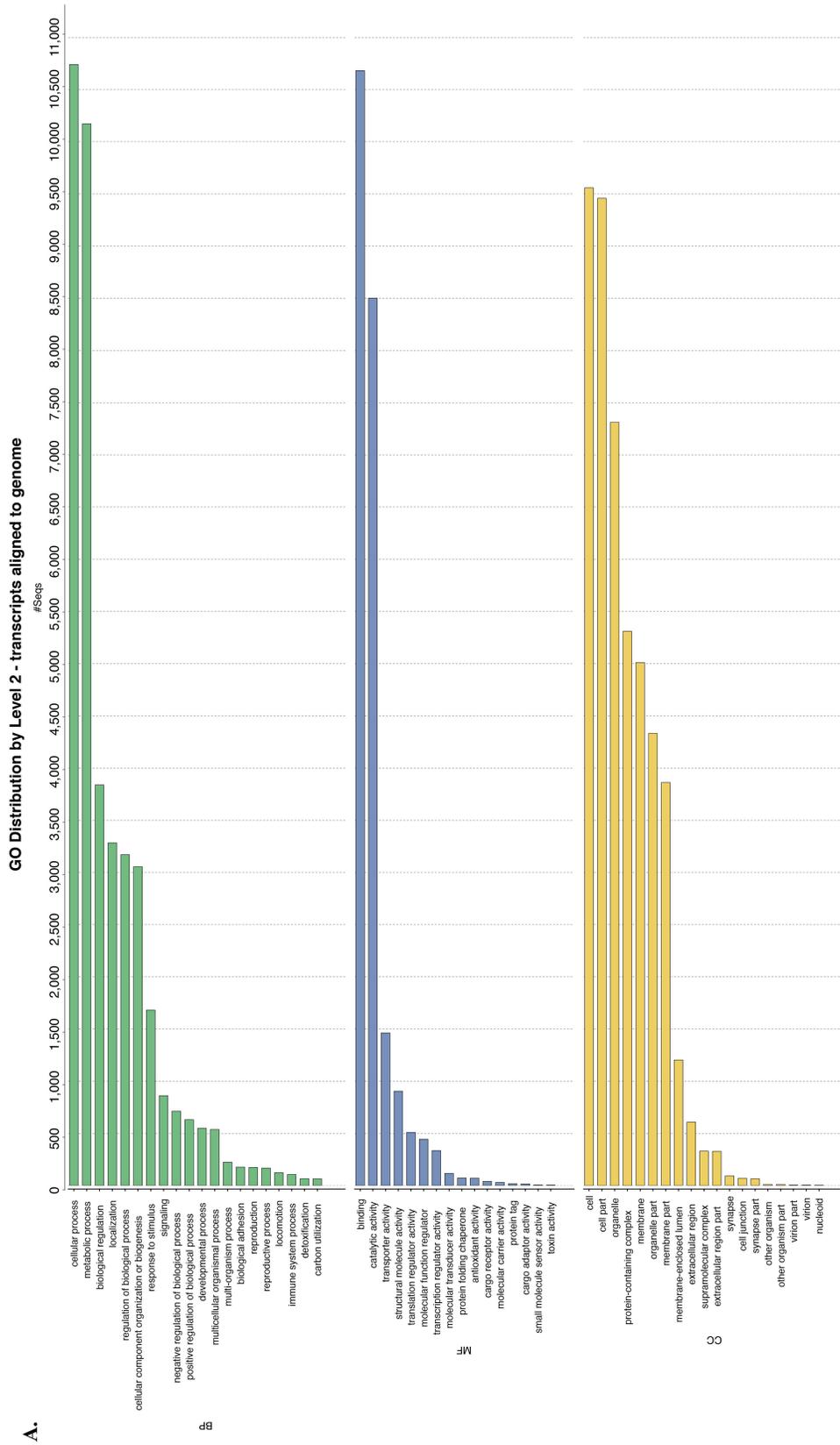
Rank	GO ID	GO Term	Number of Genes	Fisher p-value
1	GO:0005576	Extracellular region	190	2.50E-05
2	GO:0022625	Cytosolic large ribosomal subunit	36	0.0037
3	GO:0043228	Non-membrane-bounded organelle	730	0.0065
4	GO:0043232	Intracellular non-membrane-bounded organelle	730	0.0065
5	GO:0009332	Glutamate-tRNA ligase complex	1	0.0088
6	GO:0031012	Extracellular matrix	17	0.0094
7	GO:0044421	Extracellular region part	98	0.0101
8	GO:0022626	Cytosolic ribosome	59	0.0145
9	GO:0015934	Large ribosomal subunit	60	0.0152
10	GO:0005856	Cytoskeleton	175	0.0173
11	GO:0005840	Ribosome	260	0.0239
12	GO:0017102	Methionyl glutamyl tRNA synthetase complex	3	0.0262
13	GO:0070603	SWI/SNF superfamily-type complex	33	0.0335
14	GO:0005615	Extracellular space	82	0.0344
15	GO:0017101	Aminoacyl-tRNA synthetase multienzyme complex	4	0.0348

**F.**

Rank	GO ID	GO Term	Number of Gene	Fisher p-value
1	GO:0030838	Positive regulation of actin filament polymerization	13	0.00016
2	GO:1902905	Positive regulation of supramolecular fiber organization	13	0.00016
3	GO:0032273	Positive regulation of protein polymerization	14	0.0002
4	GO:0051495	Positive regulation of cytoskeleton organization	15	0.00025
5	GO:0030833	Regulation of actin filament polymerization	20	0.00061
6	GO:0031334	Positive regulation of protein complex assembly	20	0.00061
7	GO:0032271	Regulation of protein polymerization	21	0.00071
8	GO:0008064	Regulation of actin polymerization or depolymerization	22	0.00081
9	GO:0030041	Actin filament polymerization	22	0.00081
10	GO:0030832	Regulation of actin filament length	22	0.00081
11	GO:0044089	Positive regulation of cellular component biogenesis	23	0.00093
12	GO:0010638	Positive regulation of organelle organization	24	0.00106
13	GO:1902903	Regulation of supramolecular fiber organization	24	0.00106
14	GO:0008154	Actin polymerization or depolymerization	25	0.00119
15	GO:0032956	Regulation of actin cytoskeleton organization	31	0.00225

**Table 3-7.** Two genes with constitutively colony-specific splicing patterns. The putative functions are beta-1,3-glucan recognition protein (PBDG00000010286) and peritrophic membrane protein (PBDG00000008671).

Gene ID	Exon ID	Length (nt)	eCry3.1Ab1 fed 12 hours		eCry3.1Ab1 fed 24 hours		Isoline fed 12 hours		Isoline fed 24 hours	
			padj	log2FC(S/R)	padj	log2FC(S/R)	padj	log2FC(S/R)	padj	log2FC(S/R)
PBDG00000010286	E001	202	0.04061889	3.0472018	1.32E-08	2.633665432	2.09E-07	2.31017913	1.35E-07	4.12274549
	E003	106	0.4257549	0.4933681	8.71E-01	0.4071088	1.85E-15	0.27343905	1.36E-08	0.23625608
PBDG00000008671	E004	60	4.2253E-05	-2.12117479	0.029031	-2.40712866	2.53E-06	-2.3224619	2.95E-06	-2.32625442
	E005	78	0.00019877	-1.8453358	0.001742	-2.03113819	9.19E-05	-2.731343	0.000337	-2.17961178



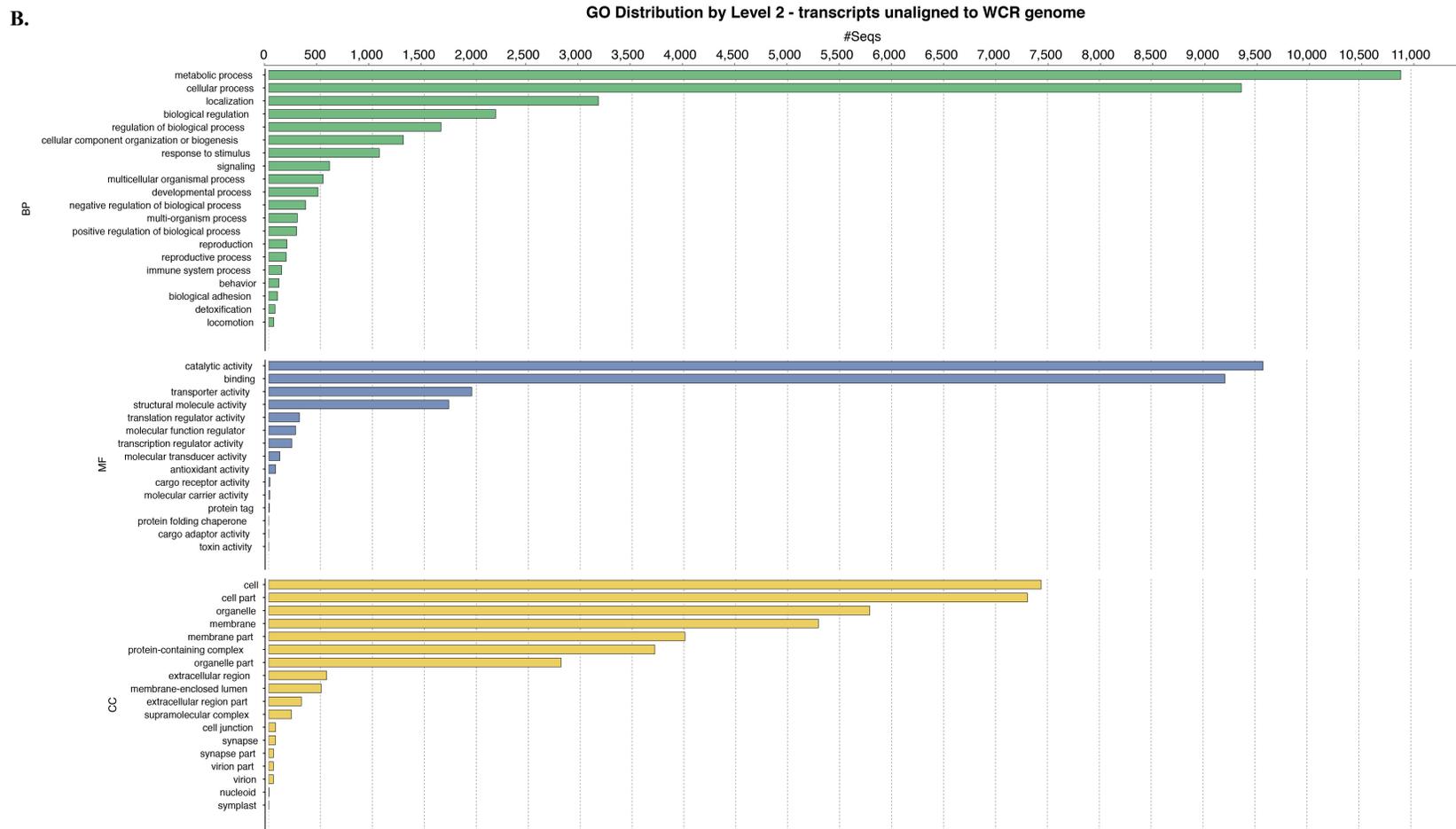
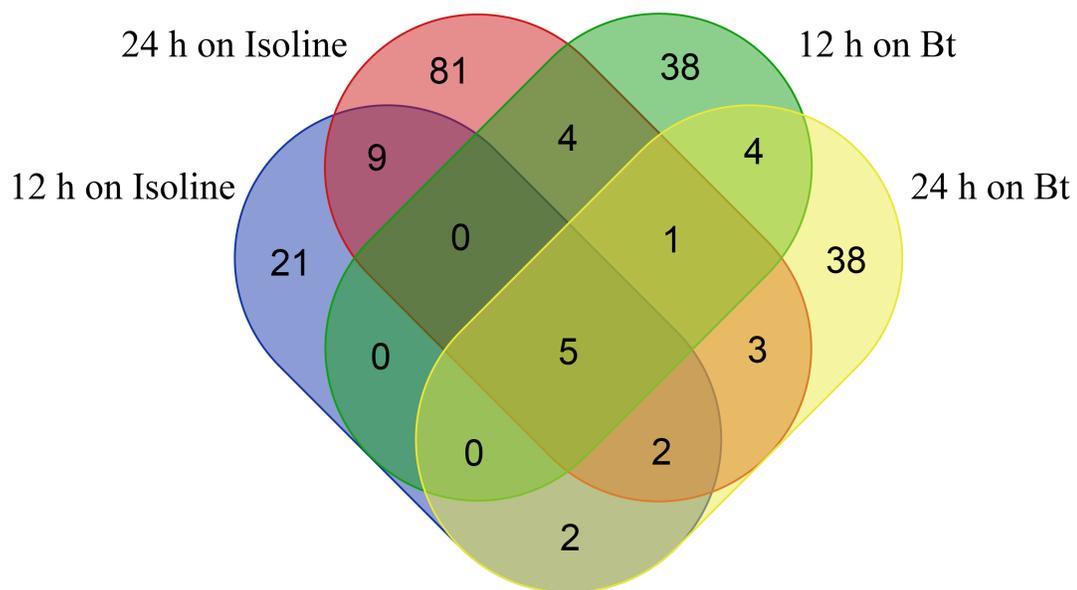
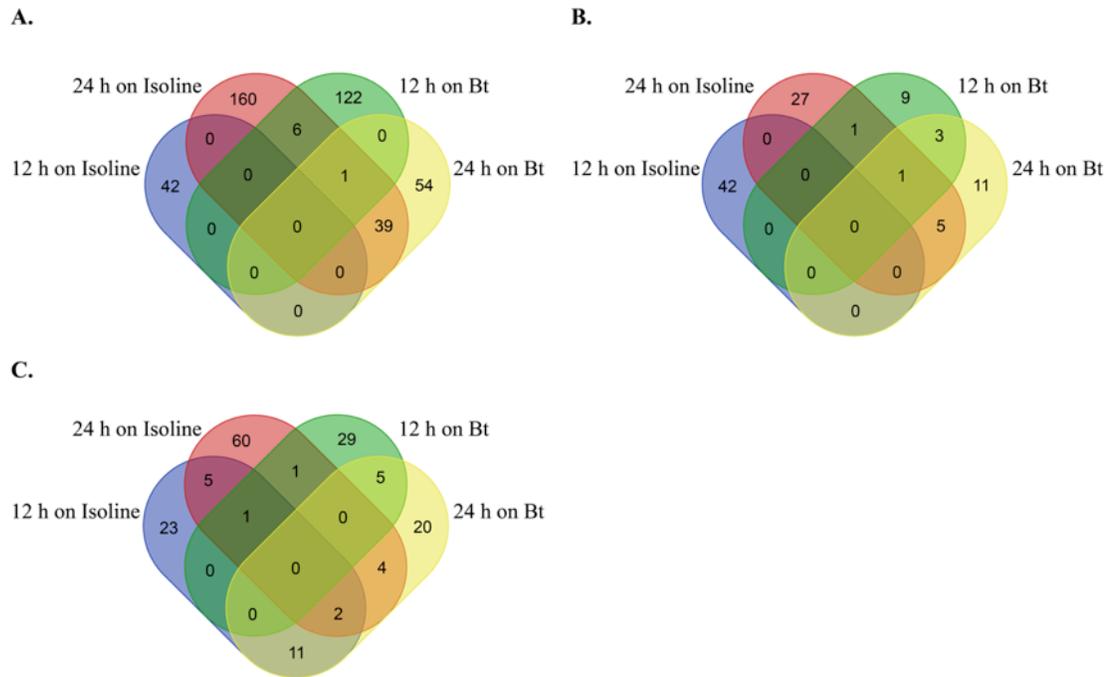


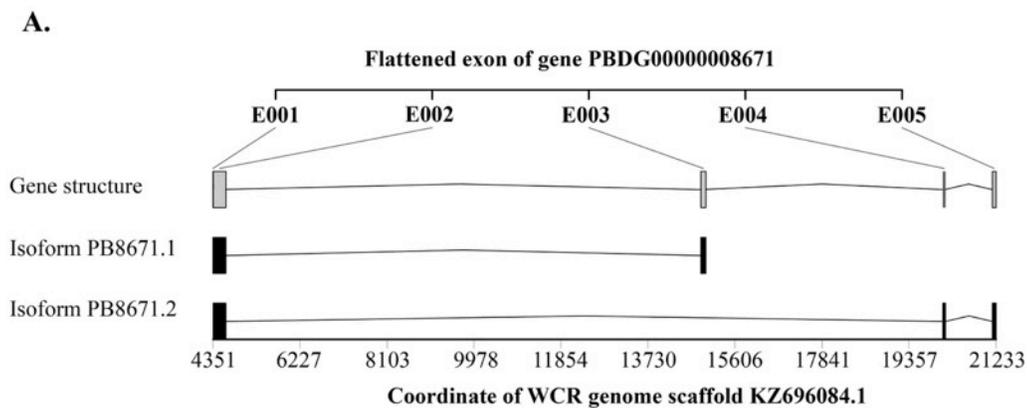
Figure 3-1. Distribution of level-2 annotated GO terms of transcripts that aligned (A) and unaligned (B) to reference genome. The GO terms are classified into biological process (BP), molecular function (MF), and cellular component (CC).



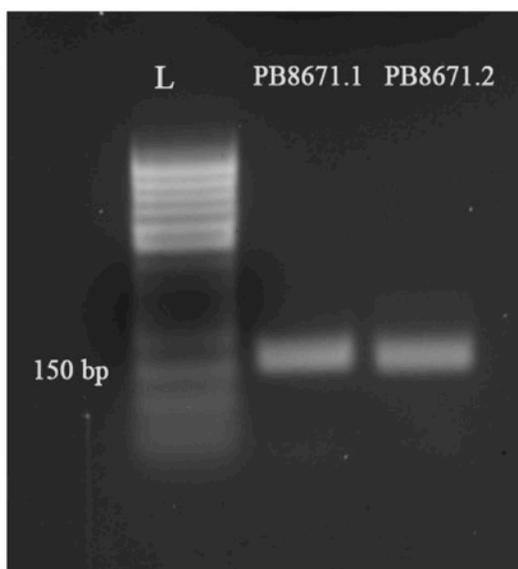
**Figure 3-2.** Venn diagram of genes with colony-specific DEUs under various experimental conditions. Isoline: feeding on non-transgenic isoline maize; Bt: feeding on eCry3.1Ab-maize diet; 12 h and 24 h are feeding duration (hours).



**Figure 3-3.** Venn diagram of enriched GO terms under different condition. The enriched GO terms are categorized into (A) Biological process (BP), (B) cellular component (CC), and (C) molecular function (MF).



**B.**



**Figure 3-4.** (A) Diagrammatic representation of gene PBDG00000008671 with its two isoforms. The gene is located on the negative strand of scaffold KZ696084.1. Alternative promoters were found for two isoforms. (B) The electrophoresis image of PCR, confirmed that exon E003 is skipped from isoform PB8671.2. The primer sets are located on exon E005, E003 (forward) and exon E002 (reverse).

```

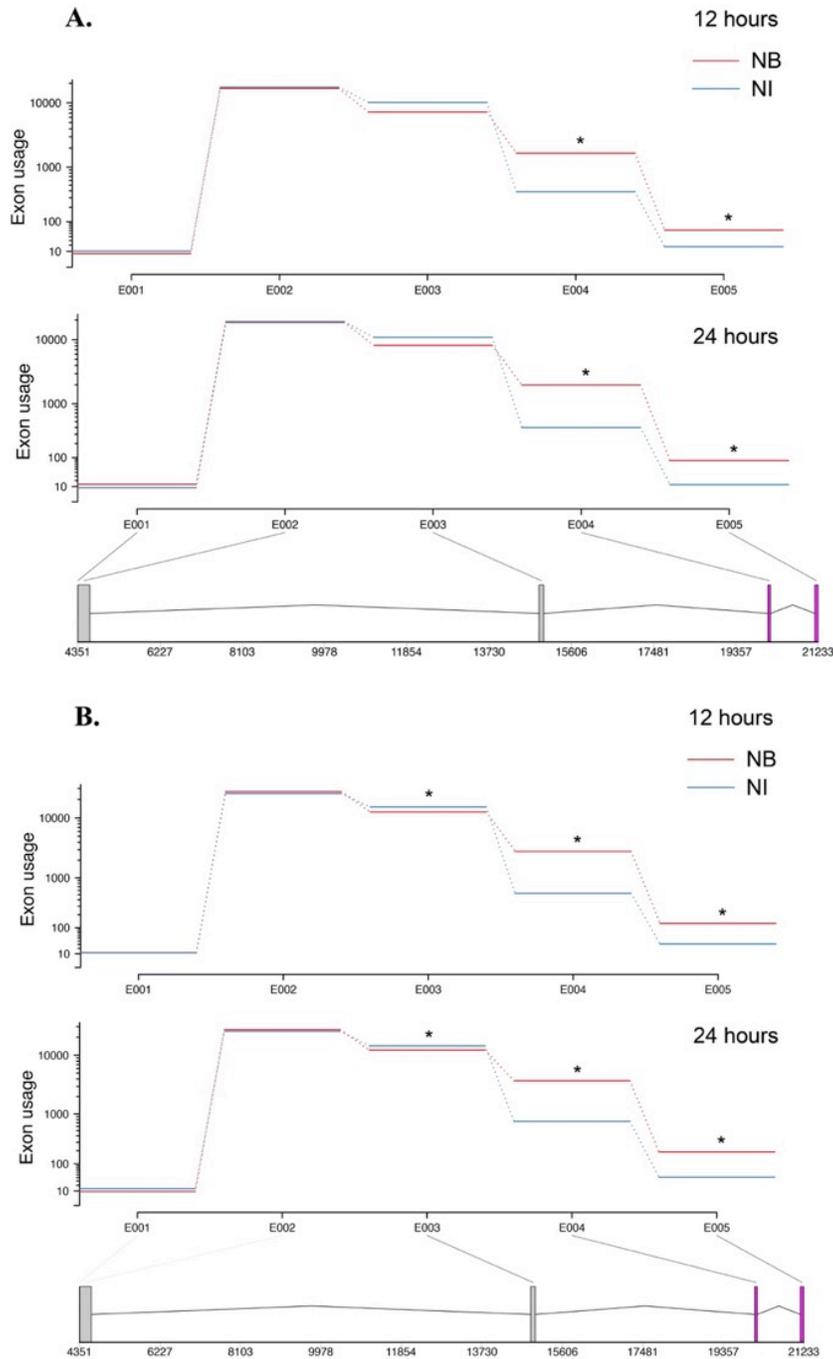
DvPMP      61 ANNTTTLAPENGTAICLQN--NVVCPAID--GPDSVFASVSDCTKFCLCSNGVPYLYNCP
PB8671.1   1  -YETAFOYP-----VTCPPFDPSAPPVYEPDPYYCDIFYECSGSTPFQFHCA
PB8671.2   1  -QEDTTLEPVEPTAAPTHRPSVTCPPFDPSAPPVYEPDPYYCDIFYECSGSTPFQFHCA

DvPMP      117 GGLHFNPSLNVCDWPENAACTNNNNDTTSTEATVTTEGTNADETTEISTTESNAAVNGT
PB8671.1   47 PGTNFDIKTNLCEF--NVDCGDR-----STVGTTTEGTOPPKF-----
PB8671.2   60 PGTNFDIKTNLCEF--NVDCGDR-----STVGTTTEGTOPPKF-----

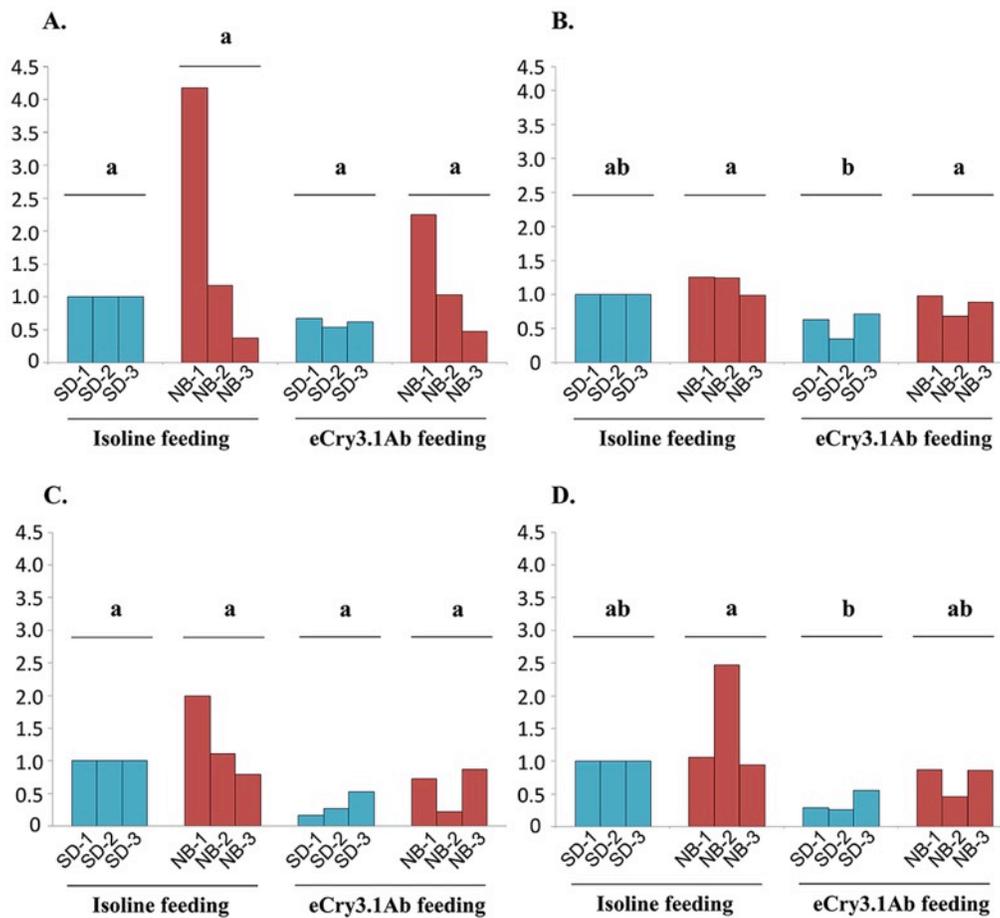
```

**Figure 3-5.** Sequencing alignment of predicted protein sequences from two isoforms and a predicted WCR peritrophic matrix protein C4 (DvPMP, GeneBank Accession number: MH256860.1). The conserved 6-cystein residues of ChtBD2-type chitin binding domain are boxed. The signal peptide of each sequences were prior estimated and removed.





**Figure 3-7.** The colony-specific differential exon usage of the novel peritrophic matrix protein gene PBDG00000008671 (alias gene ID: PB8671). Exons E004 and E005 have significantly higher usages in NB colonies than NI colony no matter feeding on eCry3.1Ab maize (A), or isoline maize (B) ( $p$ -value < 0.1).



**Figure 3-8.** The qRT-PCR quantification of novel peritrophic matrix protein PB8671 and isoform PB8671.2 in eCry3.1Ab-susceptible SD and resistance NB colonies. The relative quantification of (A) PB8671.2 only at 12 hours, (B) PB8671 in total at 12 hours, (C) PB8671.2 only at 24 hours, and (D) PB8671 in total at 24 hours were calculated by  $2^{-\Delta\Delta CT}$  method using samples of SD feeding on isoline maize as control group. SD-1, 2, 3 and NB-1, 2, 3 represent three cohorts (biological replications) of each colony. ANOVA and multiple comparison tests using Tukey adjustment were implemented on original  $\Delta Ct$  value of each sample (p-value < 0.05).

## Biolography

Abdel-Banat, B. M. A. and Koga, D. (2002) Alternative splicing of the primary transcript generates heterogeneity within the products of the gene for *Bombyx mori* chitinase. *Journal of Biological Chemistry* **277**: 30524-30534.

Alexa, A. and Rahnenfuhrer, J. (2010) topGO: enrichment analysis for gene ontology. *R package version 3.4.1* **2**: 2010.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., *et al.* (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**: 420-423.

Anders, S., Reyes, A. and Huber, W. (2012) Detecting differential usage of exons from RNA-seq data. *Genome Research* **22**: 2008-2017.

Bravo, A., Gill, S. S. and Soberón, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49**: 423-435.

Bravo, A. and Soberón, M. (2008) How to cope with insect resistance to Bt toxins? *Trends in Biotechnology* **26**: 573-579.

Campbell, P. M., Cao, A. T., Hines, E. R., East, P. D. and Gordon, K. H. J. (2008) Proteomic analysis of the peritrophic matrix from the gut of the caterpillar, *Helicoverpa armigera*. *Insect Biochemistry and Molecular Biology* **38**: 950-958.

Coates, B. S., Alves, A. P., Wang, H., Walden, K. K., French, B. W., Miller, N. J., *et al.* (2012) Distribution of genes and repetitive elements in the *Diabrotica virgifera virgifera* genome estimated using BAC sequencing. *Journal of Biomedicine and Biotechnology* **2012**.

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

Deitloff, J., Dunbar, M. W., Ingber, D. A., Hibbard, B. E. and Gassmann, A. J. (2016) Effects of refuges on the evolution of resistance to transgenic corn by the western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *Pest Management Science* **72**: 190-198.

Dong, K. (2007) Insect sodium channels and insecticide resistance. *Invertebrate Neuroscience* **7**: 17.

Fabrick, J. A., Ponnuraj, J., Singh, A., Tanwar, R. K., Unnithan, G. C., Yelich, A. J., *et al.* (2014) Alternative splicing and highly variable cadherin transcripts associated with field-evolved resistance of pink bollworm to Bt cotton in India. *PLOS ONE* **9**: e97900.

- Flagel, L. E., Swarup, S., Chen, M., Bauer, C., Wanjugi, H., Carroll, M., *et al.* (2015) Genetic markers for western corn rootworm resistance to Bt toxin. *G3: Genes|Genomes|Genetics* **5**: 399-405.
- Frank, D. L., Zukoff, A., Barry, J., Higdon, M. L. and Hibbard, B. E. (2013) Development of resistance to eCry3.1Ab-expressing transgenic maize in a laboratory-selected population of western corn rootworm (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **106**: 2506-2513.
- Gassmann, A. J., Petzold-Maxwell, J. L., Keweshan, R. S. and Dunbar, M. W. (2011) Field-evolved resistance to Bt maize by western corn rootworm. *PLoS One* **6**: e22629.
- Gonzalez-Garay, M. L. (2016) Introduction to isoform sequencing using Pacific Biosciences technology (Iso-seq). In: *Transcriptomics and Gene Regulation* (Wu, J, ed. 10.1007/978-94-017-7450-5\_6, pp. 141-160. Springer Netherlands, Dordrecht.
- Heckel, D. G., Gahan, L. J., Baxter, S. W., Zhao, J.-Z., Shelton, A. M., Gould, F., *et al.* (2007) The diversity of Bt resistance genes in species of Lepidoptera. *Journal of Invertebrate Pathology* **95**: 192-197.
- Hu, X., Chen, L., Xiang, X., Yang, R., Yu, S. and Wu, X. (2012) Proteomic analysis of peritrophic membrane (PM) from the midgut of fifth-instar larvae, *Bombyx mori*. *Molecular Biology Reports* **39**: 3427-3434.
- Huntley, R. P., Sawford, T., Mutowo-Meullenet, P., Shypitsyna, A., Bonilla, C., Martin, M. J., *et al.* (2014) The GOA database: Gene Ontology annotation updates for 2015. *Nucleic Acids Research* **43**: D1057-D1063.
- Jarosch, A., Stolle, E., Crewe, R. M. and Moritz, R. F. A. (2011) Alternative splicing of a single transcription factor drives selfish reproductive behavior in honeybee workers (*Apis mellifera*). *Proceedings of the National Academy of Sciences* **108**: 15282-15287.
- Kasai, S., Komagata, O., Okamura, Y. and Tomita, T. (2009) Alternative splicing and developmental regulation of glutathione transferases in *Culex quinquefasciatus* Say. *Pesticide Biochemistry and Physiology* **94**: 21-29.
- Kim, N., Alekseyenko, A. V., Roy, M. and Lee, C. (2006) The ASAP II database: analysis and comparative genomics of alternative splicing in 15 animal species. *Nucleic Acids Research* **35**: D93-D98.
- Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **25**: 402-408.
- Ludwick, D. C., Meihls, L. N., Ostlie, K. R., Potter, B. D., French, L. and Hibbard, B. E. (2017) Minnesota field population of western corn rootworm (Coleoptera: Chrysomelidae) shows incomplete resistance to Cry34Ab1/Cry35Ab1 and Cry3Bb1. *Journal of Applied Entomology* **141**: 28-40.

- Mathew, L. G., Ponnuraj, J., Mallappa, B., Chowdary, L. R., Zhang, J., Tay, W. T., *et al.* (2018) ABC transporter mis-splicing associated with resistance to Bt toxin Cry2Ab in laboratory- and field-selected pink bollworm. *Scientific Reports* **8**: 13531.
- Meihls, L. N., Higdon, M. L., Ellersieck, M. and Hibbard, B. E. (2011) Selection for resistance to mCry3A-expressing transgenic corn in western corn rootworm. *Journal of Economic Entomology* **104**: 1045-1054.
- Meihls, L. N., Higdon, M. L., Siegfried, B. D., Miller, N. J., Sappington, T. W., Ellersieck, M. R., *et al.* (2008) Increased survival of western corn rootworm on transgenic corn within three generations of on-plant greenhouse selection. *Proceedings of the National Academy of Sciences* **105**: 19177-19182.
- Pan, Q., Shai, O., Lee, L. J., Frey, B. J. and Blencowe, B. J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* **40**: 1413-1415.
- Rault, L. C., Siegfried, B. D., Gassmann, A. J., Wang, H., Brewer, G. J. and Miller, N. J. (2018) Investigation of Cry3Bb1 resistance and intoxication in western corn rootworm by RNA sequencing. *Journal of Applied Entomology* **142**: 921-936.
- Rees, J. S., Jarrett, P. and Ellar, D. J. (2009) Peritrophic membrane contribution to Bt Cry  $\delta$ -endotoxin susceptibility in Lepidoptera and the effect of Calcofluor. *Journal of Invertebrate Pathology* **100**: 139-146.
- Reyes, A., Anders, S. and Huber, W. (2012) Analyzing RNA-seq data for differential exon usage with the DEXSeq package.  
<http://www.bioconductor.org/packages//2.10/bioc/vignettes/DEXSeq/inst/doc/DEXSeq.pdf>.
- Rhoads, A. and Au, K. F. (2015) PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics* **13**: 278-289.
- Riddell, C. E., Lobaton Garces, J. D., Adams, S., Barribeau, S. M., Twell, D. and Mallon, E. B. (2014) Differential gene expression and alternative splicing in insect immune specificity. *BMC Genomics* **15**: 1031.
- Rodrigues, T. B., Khajuria, C., Wang, H., Matz, N., Cardoso, D. C., Valicente, F. H., *et al.* (2014) Validation of reference housekeeping genes for gene expression studies in western corn rootworm (*Diabrotica virgifera virgifera*). *PLOS ONE* **9**: e109825.
- Salz, H. K. (2011) Sex determination in insects: a binary decision based on alternative splicing. *Current Opinion in Genetics & Development* **21**: 395-400.
- Sayed, A., Nekl, E. R., Siqueira, H. A. A., Wang, H. C., Ffrench-Constant, R. H., Bagley, M., *et al.* (2007) A novel cadherin-like gene from western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae), larval midgut tissue. *Insect Molecular Biology* **16**: 591-600.

- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. and Zdobnov, E. M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210-3212.
- Strode, C., Wondji, C. S., David, J.-P., Hawkes, N. J., Lumjuan, N., Nelson, D. R., *et al.* (2008) Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* **38**: 113-123.
- Tetreau, G., Dittmer, N. T., Cao, X., Agrawal, S., Chen, Y.-R., Muthukrishnan, S., *et al.* (2015) Analysis of chitin-binding proteins from *Manduca sexta* provides new insights into evolution of peritrophin A-type chitin-binding domains in insects. *Insect Biochemistry and Molecular Biology* **62**: 127-141.
- Trapnell, C., Pachter, L. and Salzberg, S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105-1111.
- Valaitis, A. P. and Podgwaite, J. D. (2013) *Bacillus thuringiensis* Cry1A toxin-binding glycoconjugates present on the brush border membrane and in the peritrophic membrane of the Douglas-fir tussock moth are peritrophins. *Journal of Invertebrate Pathology* **112**: 1-8.
- Wang, H., Eyun, S.-i., Arora, K., Tan, S., Gandra, P., Moriyama, E., *et al.* (2017) Patterns of gene expression in western corn rootworm (*Diabrotica virgifera virgifera*) neonates, challenged with Cry34Ab1, Cry35Ab1 and Cry34/35Ab1, based on next-generation sequencing. *Toxins* **9**: 124.
- Wangila, D. S., Gassmann, A. J., Petzold-Maxwell, J. L., French, B. W. and Meinke, L. J. (2015) Susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to Bt corn events. *Journal of Economic Entomology* **108**: 742-751.
- Wu, T. D. and Watanabe, C. K. (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**: 1859-1875.
- Xiao, Y., Zhang, T., Liu, C., Heckel, D. G., Li, X., Tabashnik, B. E., *et al.* (2014) Mis-splicing of the ABCC2 gene linked with Bt toxin resistance in *Helicoverpa armigera*. *Scientific Reports* **4**: 6184.
- Zhao, Z., Meihls, L. N., Hibbard, B. E., Ji, T., Elsik, C. G. and Shelby, K. S. (2019) Differential gene expression in response to eCry3.1Ab ingestion in an unselected and eCry3.1Ab-selected western corn rootworm (*Diabrotica virgifera virgifera* LeConte) population. *Scientific Reports* **9**: 4896.
- Zukoff, S. N., Ostlie, K. R., Potter, B., Meihls, L. N., Zukoff, A. L., French, L., *et al.* (2016) Multiple assays indicate varying levels of cross resistance in Cry3Bb1-selected field populations of the western corn rootworm to mCry3A, eCry3.1Ab, and Cry34/35Ab1. *Journal of Economic Entomology* **109**: 1387-1398.



## **Chapter 4**

Preliminary analysis of genomic variations in association with eCry3.1Ab resistance in western corn rootworm

## Abstract

Single nucleotide polymorphisms (SNPs) and small insertions and deletions (Indels) are highly abundant in insect genomes. They can provide useful information on population genetics, mapping of economically important traits, and resistance monitoring in insect resistant management (IRM) programs. Mutations in receptors as well as other related proteins may cause resistance to Bt and other insecticides. To understand the resistance mechanism against eCry3.1Ab and provide a marker set for further genetics studies of the western corn rootworm, we re-analyzed the genetic variations of eCry3.1Ab-resistant and susceptible western corn rootworm colonies from previously reported RNA-seq and PacBio transcriptome Iso-seq data. We were able to identify a set of polymorphic genomic variants, including SNPs and indels, which distinguish the lab-selected eCry3.1Ab-resistant colony from its counterpart unselected, eCry3.1Ab-sensitive colony. The genetic diversity within the resistant colony is reduced, based on principal component analysis (PCA). We also identified a SNP at a putative ABC transporter causing a nonsynonymous amino acid substitution. Our results provide preliminary analysis of genomic variations in western corn rootworms and further studies will facilitate the understanding of the genetic bases of eCry3.1Ab resistance.

## Introduction

In agricultural ecosystems, insects experience strong selection pressure from management strategies. This selection pressure allows insects to develop resistance resulting in failure to control the insects in the field. *Bacillus thuringiensis* (Bt) toxins Cry3Bb1, mCry3A, eCry3.1Ab, and Cry34/35Ab1 have been commercialized for management of the western corn rootworm (WCR, *Diabrotica virgifera virgifera* LeConte) and show variable effectiveness (Ludwick and Hibbard 2016). Since 2003 with the commercialization of transgenic maize hybrids expressing the Cry3Bb1 protein, these toxins have been widely adopted in the US Corn Belt as the primary rootworm management approach (Vaughn et al. 2005). Field-evolved resistance to Cry3Bb1 maize has been independently reported at multiple locations (Gassmann et al. 2011; Ludwick et al. 2017; Wangila et al. 2015). Resistance to mCry3A (Gassmann et al. 2014) and to Cry34/35Ab1 (Gassmann et al. 2016; Ludwick et al. 2017) have also been identified in the corn field. Cross-resistance among Cry3Bb1, mCry3A and eCry3.1Ab has been demonstrated in some WCR populations (Gassmann et al. 2014; Zukoff et al. 2016).

Both field surveys and laboratory-based selection studies (Frank et al. 2013; Meihls et al. 2011; Meihls et al. 2012; Meihls et al. 2008) have documented that WCR can rapidly develop resistance to Bt, indicating that the resistance alleles to Bt toxins are not rare. All four rootworm-targeting Bt toxins are low to moderate dose (Alstad and Andow 1995; Andow et al. 2015). Resistance to Bt does not compromise fitness in WCR, suggesting that new insect resistance management (IRM) strategies other than traditional “high-dose” structured refuges are necessary for rootworm-targeting Bt-maize (Andow et al. 2015).

One of the critical steps for implementing IRM is resistance monitoring. Current techniques are based on bioassays, which rely on phenotypic responses of insect larvae and their progenies to Bt (Huang 2006). Molecular marker-based approaches could directly detect the resistant alleles and distinguish the heterozygous and homozygous individuals without bioassays or crossing experiments, which could reduce the cost of greenhouse procedures such as colony maintenance and bioassays. Genetic markers would also be used in early prediction of resistance emergence and control failure. However, it requires having markers that are linked to resistance genes, which are still not clear for eCry3.1Ab in WCR.

Molecular markers also provide powerful tools to understand the genetic basis of biological processes, adaptation to the environment, and distribution of insects. In WCR, microsatellites (SSR) and amplified fragment length polymorphism (AFLP) markers have been used to study the population structures, introduction to Europe, and adaptation to maize-soybean rotation (Kim and Sappington 2005a; Kim et al. 2008; Miller et al. 2007; Miller et al. 2005; Miller et al. 2009). Single nucleotide polymorphisms (SNP) are highly abundant in genomes and can improve the quality of genetic mapping. Linkage mapping using SNP markers has been used to identify QTLs that link to organophosphate insecticide and Cry3Bb1 resistance (Coates et al. 2016; Flagel et al. 2015).

The protein eCry3.1Ab is an engineered Bt toxin modified from Coleoptera-targeting mCry3A with lepidoptera-target Cry1Ab sequences at its C-terminal, which increased the efficacy against WCR (Walters et al. 2010). WCR has developed resistance to eCry3.1Ab in laboratory selection (Frank et al. 2013). Although some resistant colonies to Cry3Bb1 and mCry3A show cross-resistance to eCry3.1Ab (Zukoff

et al. 2016), control failure has not been reported in the refereed literature because the eCry3.1Ab is never offered as a single trait product. Although binding assays suggest that eCry3.1Ab binds to midgut epithelial cell membrane with multiple binding sites (Walters et al. 2010), details of this interaction and resistance mechanisms are currently unknown.

High-throughput RNA sequencing (RNA-seq) technology allows detection and quantification of transcriptome-wide expression changes. It also provides a path to assess genetic variation. Compared to genome sequencing, RNA-seq focuses on the exonic regions of genes, in which the variants might directly result in changes of protein sequence and function. In this study, we surveyed genomic variations in association with eCry3.1Ab resistance using RNA-seq data. These sequence variants will be further used to study the eCry3.1Ab resistance mechanism and population genetics in WCR, and will support development of marker sets for monitoring Bt resistance in WCR.

## **Material and Methods**

### **Data source**

WCR RNA-seq data has been previously described by Zhao et al. (2019) (also see Chapter 2). The data were generated from a laboratory-selected non-diapause eCry3.1Ab-resistant WCR colony (NB) and its sibling non-selected, eCry3.1Ab sensitive NI colony. The selection processes have been described by Frank et al. (2013). Before RNA extractions, three cohorts of each colony were fed on eCry3.1Ab-expressing transgenic and isoline sprouted maize seedlings for 12 and 24 hours (Table 4-1). RNA-seq libraries were prepared separately for each of the treatments and biological replicates. Here we focus on genetic variation at the cohort level: thus within each cohort,

the RNA-seq reads of different treatments were merged to increase the sequence coverage.

The reference WCR genome assembly (version 1) was downloaded from the NCBI genome database ([https://www.ncbi.nlm.nih.gov/assembly/GCA\\_003013835.1/](https://www.ncbi.nlm.nih.gov/assembly/GCA_003013835.1/)). It is assembled from genomic sequencing data of a non-diapause inbred strain PED-12 from USDA-ARS at Brookings, SD. The basic statistics of genome assembly have been reviewed in Chapter 1. We also used the genome annotation file from previous PacBio Iso-seq (Chapter 3) to obtain gene structure information for annotating variants in this study.

### **RNA-seq reads trimming and alignment to reference genome**

To obtain “clean reads” free of adapters, low quality reads and mitochondrial sequences we adopted the same trimming procedures described in Zhao et al. (2019) on RNA-seq data. The residue of sequencing adapters was first removed by FastqMcf (version 1.04.8, <https://github.com/ExpressionAnalysis/ea-utils/blob/wiki/FastqMcf.md>). Trimmomatic (version 0.36) (Bolger et al. 2014) was employed to remove low quality reads. As the final step of trimming we aligned reads to the WCR mitochondrial genome (Coates 2014) using Bowtie 2 (version 4.7.7) (Langmead and Salzberg 2012) and discarded all the concordantly aligned reads.

The clean reads were pooled by cohort and then aligned to the reference WCR genome by HISAT2 (version 2.0.1), a fast spliced aligner that uses Burrow-wheeler transform-based mapping algorithms (Kim et al. 2015). The splice feature of HISAT2 allows RNA-seq reads to cover splice junctions between exon and intron. The output alignment files were in SAM format and converted to binary BAM format using

SAMtools (Li et al. 2009). We further sorted, added read groups, and marked duplicated reads of alignment files using Picard (version 2.4.1, <http://broadinstitute.github.io/picard/>).

### **Variant calling, filtering and principal component analysis**

The SNP as well as small insertions and deletions (Indels) were called from alignment files using GATK (version 3.5) best practices pipeline for variant calling on RNA-Seq data (DePristo et al. 2011). We first used the SplitNCigarReads package from GATK to split reads that expand over exon boundaries and reassigned the mapping quality. Then, we used GATK HaplotypeCaller to call SNPs and Indels from each dataset and stored the output as a variant call format (VCF) file.

To reduce the false positives, it was necessary to apply filters to genomic variations. However, the variant quality score recalibration (VQSR) is not applicable to RNA-seq based variants. We can only use hard filters. The arbitrary hard filters set includes the following criteria: (1) clustering variants with 3 variants within 35-bases window, (2) variants with Fisher Strand value greater than 30, and (3) variants with quality by depth value less than 2.0 and (4) variants with minimum coverage depth less than 25 at the locus. Variants that met any of above criteria will be discarded. The variants with the “PASS” flag are considered as high quality in this experiment. This resulted in preliminary genomic variant sets of NB and NI colonies.

To characterize the genetic diversity and population structure among cohorts and colonies, we applied principal component analysis (PCA) to SNP sets of three NB and three NI colonies. The GDS file for PCA analysis was created from a merged VCF file

containing variants from six samples. PCA analysis was conducted using the R package SNPRelate (Zheng et al. 2012) with default settings.

### **Variant annotation and intersection analysis**

We further used BCFtools (version 1.9) (Li 2011) to annotate the variants using PacBio Iso-seq transcriptome (see Chapter 3). The Iso-seq post-analysis script Cupcake ToFU generates a GFF file containing alignment of transcripts to the WCR genome. The output GFF file was converted to standard GFF3 format by a custom Perl script and further converted to bed format by BEDOPS (version 2.4.36) (Neph et al. 2012). The annotation script from BCFtools was employed to map transcripts from the PacBio Iso-seq study to gene loci containing genomic variants. Functional annotations of those transcripts were obtained by BLASTX against the NCBI non-redundant protein database (nr) using Blast2GO (Götz et al. 2008). To detect the possible markers that separate NB and NI populations, we also used BCFtools intersection script to extract NB-specific variants that were present in all three NB cohorts but none of the NI cohort and *vice versa*.

Flagel et al. (2015) reported that in WCR, a section of 115-135 cM from linkage group LG8 are statistically associated with Cry3Bb1 resistance using linkage mapping analysis. There are 39 SNPs found in that section, including two most significant SNPs (CRW1683 and CRW1892) and others with putative Bt resistance function, i.e. ABC transporter subfamily B (CRW424) and cytochrome P450 (CRW918). To test if the same variants existed in NB and NI colonies, we first mapped the flanking regions of those SNPs to the WCR genome using GMAP (version 2014-08-19) (Wu and Watanabe 2005) and searched all NB or NI specific variants from this region.

## Results

### Variant calling and PCA analysis

From the RNA-seq results of three cohorts of NB and NI we are able to detect 64,613 to 76,837 filtered variants from each cohort using HISAT2 alignment and GATK variant discovery tools. Most of the variants were SNPs; only a small portion of each variant set was indels. Among the 33,721 SNPs and 1,436 indels were nonpolymorphic variants that were shared by all 6 samples. Using the current genome annotation obtained from PacBio Iso-seq, about 60% of variants were located in annotated regions, whereas about 50% of aligned genes from Iso-seq have SNPs or indels (Table 4-1).

The PCA on SNP sets of three NB and three NI cohorts obtained two components with eigenvalues larger than 1, which accounted for 34.58% and 20.01% of total variation (Table 4-2). The PCA plot of eigenvector 1 and 2 demonstrated that NB and NI colonies can be separated by eigenvector 1. However, eigenvector 2 could not separate NI and NB colonies. Three NB cohorts were clustered, suggesting that the NB cohorts are from a genetically distinct population. The NI cohorts were less clustered, indicating a higher genetic diversity in the NI population (Figure 4-1).

### Variant intersection, annotation and functional analysis

To select a list of highly stringent genomic variants that could separate NB from NI colony, we used the intersection script from BCFTools package to extract variants that were present in all NB cohorts but in none of NI cohorts, and *vice versa*. We identified 1,836 SNPs and 98 indels that were private to NB cohorts, among which 1,196 variants were located in 619 gene loci that were already annotated from the Iso-seq genomic feature data. In contrast, NI cohorts had 2,184 SNPs and 109 indels. Among them 1,502

variants were located in 713 gene loci. There were 513 and 608 gene loci in the WCR genome containing only NB or NI specific variants, respectively, which suggested some gene loci host more than one colony specific variant. There were 105 gene loci containing both NB and NI specific variants, suggesting those genes have both NB and NI specific alleles (Figure 4-2).

It has been well acknowledged that cadherin, aminopeptidase N (APN), alkaline phosphatase (ALP) and ABC transporters can be Bt receptors and their functions might be critical to Bt toxicity (Bravo et al. 2007; Heckel 2012; Pardo-López et al. 2013). We identified an ABC transporter subfamily A (PB.7654) and an ABC transporter family C gene (PB.5841) with NB specific SNPs. Sequence analysis demonstrated that the SNP in ABCA transporter was located in 3'-UTR region, while one SNP of ABCC transporter caused amino acid substitution from arginine to lysine (Arg1180Lys) (Table 4-3). A motif search by Pfam (Sonnhammer et al. 1998) showed the substitution of ABCC transporter to be located in an ATP-binding domain (Pfam: PF00005) at the second nucleotide-binding domain (NBD), in between the conserved Walker A and Walker B nucleotide binding motif (Figure 4-3). The functional alteration of this substitution, as well as this specific ABCC transporter has not been studied. We also identified a cadherin-like protein (PB.112) hosting a NB specific SNP. However, these SNPs caused a synonymous nucleotide substitution that does not alter the amino acid sequence (Table 4-3).

We tested 39 SNPs markers that were located in linkage group LG8. One NB-specific SNP (SNP id: KZ708623.1\_20125\_T\_C) detected in the WCR genome was identical to the marker CRW918. The gene with that SNP had a predicted function of Cytochrome P450 6k1 (Flagel et al. 2015). Sequence analysis showed a substitution of

isoleucine to valine in the predicted protein sequence. However, we could not obtain the full open reading frame (ORF) due to the lack of genomic feature information at this locus.

## Discussion

Since the first Bt-resistant insect was reported (McGaughey 1985), the identification of Bt resistance genes has been a primary focus in the efforts to decipher resistance mechanisms. Biochemical approaches have studied the molecular interactions during toxin activation, passage through the peritrophic membrane, receptor binding and pore formation, and have resulted in the elucidation of Bt mode of action in many Lepidopteran insects (Bravo et al. 2007; Ferré and Van Rie 2002). On the other hand, forward genetics approaches are able to map genes that control Bt resistance to the genetic maps using linkage mapping without any prior biochemical knowledge. Genetic mapping led to the first discovery of a mutation in the *Heliothis virescens* ABC transporter C2 (HvABCC2) that causes resistance to Cry1Ac (Gahan et al. 2010). The HvABCC2 had not shown any affinity to Bt in binding assays. The only prerequisites for linkage analysis are (1) a mapping population carrying Bt resistance traits and (2) a set of markers to construct linkage map. Creating a set of SNP markers will facilitate progress in understanding the genetic and molecular mechanism of eCry 3.1Ab resistance in WCR.

In addition to gene expression and alternative splicing analysis presented in Chapters 2 and 3, we further detected the genetic variations that might be associated with eCry3.1Ab resistance. Using RNA-seq data from NB (eCry3.1Ab-resistant) and NI (eCry3.1Ab-susceptible) colonies respectively, we were able to identify 61,545 to 72,942

SNPs and 3,068 to 3,895 indels of each sample respectively from the current genome assembly. Approximately 50% of SNPs and 59% of indels were polymorphic between or within colonies. We detected smaller variant numbers than a field-collected sample characterized by Flagel et al. (2014), indicating that lab-selected NB and NI colonies have less genetic diversity compared to field-resistant colonies.

Previous studies on several lepidopteran insect species reported that null mutations by insertion or deletion of cadherin (Gahan et al. 2001), APN (Zhang et al. 2009), and ABC transporters (Gahan et al. 2010; Tay et al. 2015) confer resistance to Bt toxins. In the silkworm, *Bombyx mori*, an insertion of a codon (an amino acid) at the second outer loop of ABCC4 caused resistance to Cry1Ab, suggesting that a single amino acid insertion at a critical region is sufficient to develop Bt resistance (Atsumi et al. 2012). We reported an NB-specific SNP in a putative ABCC transporter that caused the substitution of arginine to lysine at residue 1140. This gene appeared to be a fully functional ABC transporter with two nucleotide-binding domains (NBDs), two transmembrane domains (TMDs), and a signature sequence LSGGD. The substitution was located in NBD-2 domain (Figure 4-3). The possible function of this ABC transporter, as well as the effect of the identified amino acid substitution, is yet to be studied.

Genome-wide SNPs provide a powerful measurement of population structure. The ideal approach would be to genotype a large number of individuals per population. In the present study the number of samples was limited due to the initial experimental design. Each RNA-seq sequencing library already consisted of multiple individuals from a cohort each of the NB or NI colony. In this case, PCA analysis on polymorphic SNP markers

still showed clear separation between NB and NI cohorts. Both colonies were derived from the same initial WCR population. The initial genetic diversity within the population was increased by reciprocally backcrossing of lab-reared USDA-ARS non-diapausing colony (Brookings) with a WCR feral population (Frank et al. 2013). The initial selection pressure imposed by eCry3.1Ab-expressing maize dramatically reduced the number of individuals, which provided a strong bottleneck effect. Also, as laboratory selection starts with a relatively smaller population, the founder effect would further reduce the genetic diversity. This process is suggested in PCA plots that even with limited sample numbers, the NI colony still had larger variation among cohorts than NB. A similar result was reported from a field-collected Cry3Bb1-resistant colony Hopkinton; SNP markers developed from Hopkinton separated it from non-selected WCR (Brookings). But the same marker set did not separate other Cry3Bb1-resistant colonies from their susceptible counterparts (Rault et al. 2018).

The associated markers with eCry3.1Ab resistance will greatly reduce the efforts and costs in monitoring resistance in IRM program. However, whether the resistance mechanisms from lab-selected populations can be extended to field-evolved resistant populations is not known. Studies from insecticide resistance suggested lab-selection favors polygenic resistance such as metabolic adaptation to toxins, while field-selection favors monogenic resistance such as the mutation of receptors (French-Constant 2013). For the WCR, laboratory selection experiments (Deitloff et al. 2016; Frank et al. 2013; Meihls et al. 2011; Meihls et al. 2012; Meihls et al. 2008) have mimicked field resistance quite well in terms of generations of selection and resistance levels (Gassmann et al. 2011; Ludwick and Hibbard 2016; Ludwick et al. 2018; Ludwick et al. 2017; Zukoff et al.

2016). In addition to that, we were able to detect one NB-specific SNP marker located on the associated loci of Cry3Bb1 resistance, which was identified in another lab-selected Cry3Bb1-resistant colony HopR (Flagel et al. 2015). The shared identical SNP marker of two distinct colonies might indicate the possible mechanism of cross-resistance between Cry3Bb1 and eCry3.1Ab. However, the present study as well as other published WCR SNP datasets (Coates et al. 2016; Flagel et al. 2014; Flagel et al. 2015; Rault et al. 2018) are called from sequencing of gene transcripts. Whether genetic variations in introns or intergenic regions would be associated with Bt resistance have not been characterized yet.

In conclusion, with the WCR genome assembly and high-throughput sequencing technology, we conducted a preliminary analysis of genomic variations in eCry3.1Ab-resistant and susceptible colonies and identified genes with colony-specific variants, which might relate to resistance to eCry3.1Ab. However, to further test the relatedness the variants to Bt resistance we need to perform genome-wide association analysis (GWAS) on variants at the individual level of multiple populations. GWAS analysis would allow the characterization of population structures of resistant colonies, fine mapping of eCry3.1Ab resistance traits, identify resistance-related genes, and ultimately provide techniques for monitoring eCry3.1Ab resistance in the field.

**Table 4-1.** Statistics of detected genomic variants of each sample. The RNA-seq reads of each cohort were merged from all the bioassay treatments (described in Chapter 2) implemented on that cohort.

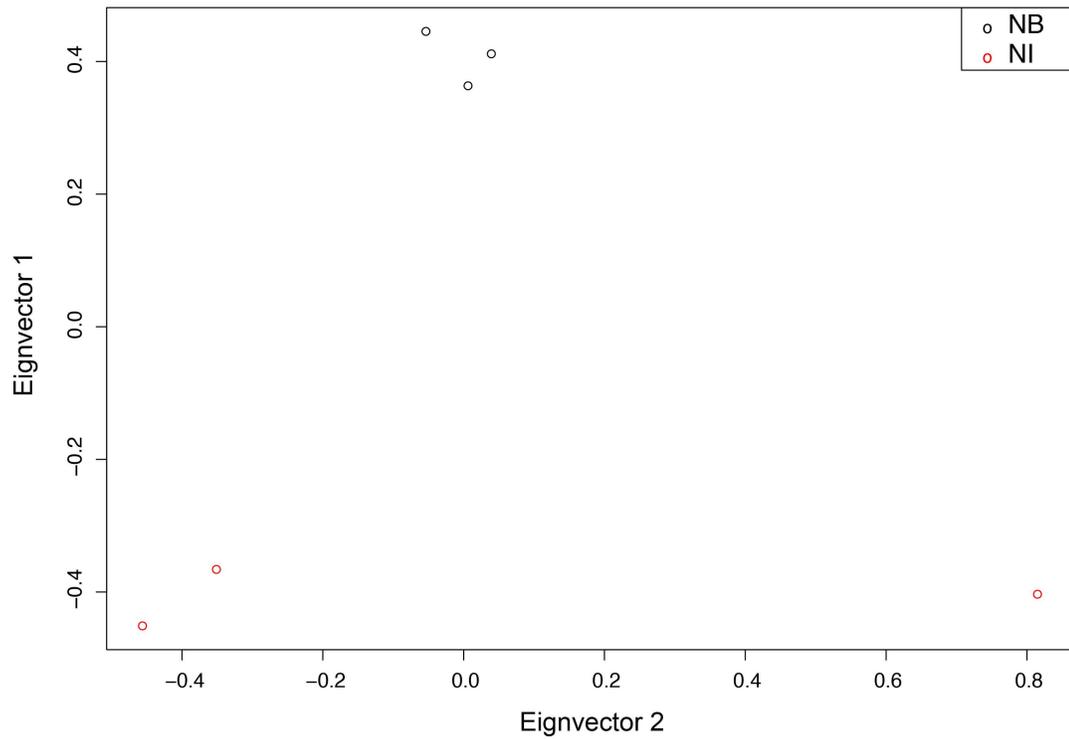
Colony	Cohort	Variant number		Number of variant w/annotation	Iso-seq genes w/variant
		SNP	Indel		
NB	R1	64,484	3,270	40,498	5,748
	R2	72,942	3,895	43,774	6,004
	R3	63,106	3,335	40,499	5,881
NI	S1	70,213	3,654	42,841	6,014
	S2	71,235	3,808	44,067	6,058
	S3	61,545	3,068	40,395	5,997

**Table 4-2.** Statistical parameters of principal component analysis (PCA).

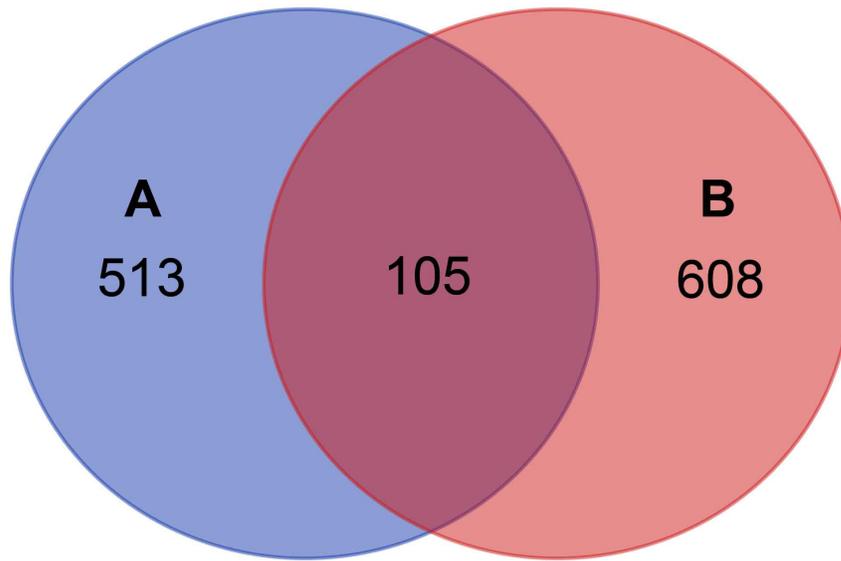
	<b>Eigenvalue</b>	<b>Variance for proportion</b>
Eigenvector 1	1.729072	34.58%
Eigenvector 2	1.000672	20.01%
Eigenvector 3	0.8669498	17.34%
Eigenvector 4	0.722153	14.44%
Eigenvector 5	0.6811536	13.62%
Eigenvector 6	-2.83107E-15	-5.66214E-16

**Table 4-3.** Pattern of NB-specific variants in potential Bt resistance associated genes

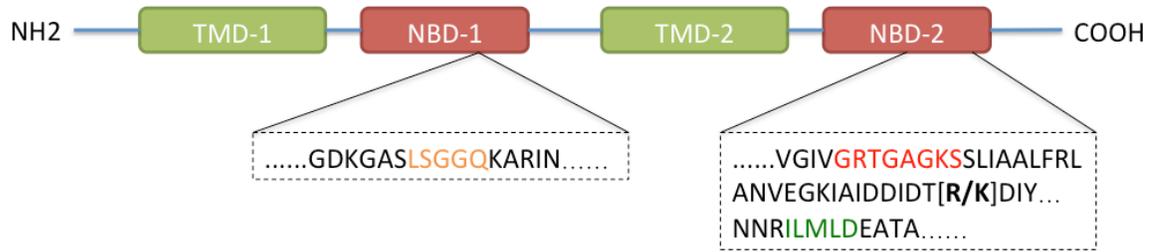
<b>Variant locus</b>	<b>Mutation</b>	<b>Gene ID</b>	<b>Functional annotation</b>	<b>Mutation type</b>
KZ688680.1: 626669	G to A	PB112	Cadherin-like protein	Same-sense
KZ691562.1: 83829	C to T			Same-sense
KZ691562.1: 109713	A to G			Same-sense
KZ691562.1: 111278	A to G			Same-sense
KZ691562.1: 111317	G to A	PB5841	ABC transporter subfamily C	Same-sense
KZ691562.1: 111349	G to A			Arg1180Lys
KZ691562.1: 111494	G to A			Same-sense
KZ691562.1: 115529	C to T			Same-sense
KZ693868.1: 39757	T to C	PB7654	ABC transporter subfamily A	3'UTR region



**Figure 4-1.** Plot of principal component analysis of three cohorts of eCry3.1Ab-selected resistant colony (NB) and unselected eCry3.1Ab-susceptible colony (NI). The eigenvectors 1 and 2 were calculated in PCA using polymorphic SNP datasets. The NB cohorts were clustered together while higher degree of variation was observed among NI cohorts.



**Figure 4-2.** Venn diagram of the number of gene loci containing NB (A) and NI (B) specific genetic variations. The information of gene loci was from genomic feature obtained by PacBio Iso-seq experiment, described in Chapter 3.



**Figure 4-3.** Sequence analysis of a putative WCR ABC transporter subfamily C protein. TMD-1 and TMD-2 are the two transmembrane domains, while NBD-1 and NBD-2 are the two nucleotide-binding domains. NBD-1 has a conserved ABC transporter signature LSGGQ motif (sequence marked in yellow). The amino acid substitution is located at NBD-2 domain, in between the conserved Walker A (sequence marked in red) and Walker B (sequence marked in green).

## Bibliography

- Alstad, D. N. and Andow, D. A. (1995) Managing the evolution of insect resistance to transgenic plants. *Science* **268**: 1894-1896.
- Andow, D. A., Pueppke, S. G., Schaafsma, A. W., Gassmann, A. J., Sappington, T. W., Meinke, L. J., *et al.* (2015) Early detection and mitigation of resistance to Bt maize by western corn rootworm (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **109**: 1-12.
- Atsumi, S., Miyamoto, K., Yamamoto, K., Narukawa, J., Kawai, S., Sezutsu, H., *et al.* (2012) Single amino acid mutation in an ATP-binding cassette transporter gene causes resistance to Bt toxin Cry1Ab in the silkworm, *Bombyx mori*. *Proceedings of the National Academy of Sciences* **109**: E1591-E1598.
- Bolger, A. M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.
- Bravo, A., Gill, S. S. and Soberón, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49**: 423-435.
- Coates, B. S. (2014) Assembly and annotation of full mitochondrial genomes for the corn rootworm species, *Diabrotica virgifera virgifera* and *Diabrotica barberi* (Insecta: Coleoptera: Chrysomelidae), using next generation sequence data. *Gene* **542**: 190-197.
- Coates, B. S., Alves, A. P., Wang, H., Zhou, X., Nowatzki, T. M., Chen, H., *et al.* (2016) Quantitative trait locus mapping and functional genomics of an organophosphate resistance trait in the western corn rootworm, *Diabrotica virgifera virgifera*. *Insect Molecular Biology* **25**: 1-15.
- Deitloff, J., Dunbar, M. W., Ingber, D. A., Hibbard, B. E. and Gassmann, A. J. (2016) Effects of refuges on the evolution of resistance to transgenic corn by the western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *Pest Management Science* **72**: 190-198.
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**: 491-498.
- Ferré, J. and Van Rie, J. (2002) Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. *Annual Review of Entomology* **47**: 501-533.
- French-Constant, R. H. (2013) The molecular genetics of insecticide resistance. *Genetics* **194**: 807-815.
- Flagel, L. E., Bansal, R., Kerstetter, R. A., Chen, M., Carroll, M., Flannagan, R., *et al.* (2014) Western corn rootworm (*Diabrotica virgifera virgifera*) transcriptome assembly and genomic analysis of population structure. *BMC Genomics* **15**: 1-13.

- Flagel, L. E., Swarup, S., Chen, M., Bauer, C., Wanjugi, H., Carroll, M., *et al.* (2015) Genetic markers for western corn rootworm resistance to Bt toxin. *G3: Genes|Genomes|Genetics* **5**: 399-405.
- Frank, D. L., Zukoff, A., Barry, J., Higdon, M. L. and Hibbard, B. E. (2013) Development of resistance to eCry3.1Ab-expressing transgenic maize in a laboratory-selected population of western corn rootworm (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **106**: 2506-2513.
- Gahan, L. J., Gould, F. and Heckel, D. G. (2001) Identification of a gene associated with Bt resistance in *Heliothis virescens*. *Science* **293**: 857.
- Gahan, L. J., Pauchet, Y., Vogel, H. and Heckel, D. G. (2010) An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLOS Genetics* **6**: e1001248.
- Gassmann, A. J., Petzold-Maxwell, J. L., Clifton, E. H., Dunbar, M. W., Hoffmann, A. M., Ingber, D. A., *et al.* (2014) Field-evolved resistance by western corn rootworm to multiple *Bacillus thuringiensis* toxins in transgenic maize. *Proceedings of the National Academy of Sciences* **111**: 5141-5146.
- Gassmann, A. J., Petzold-Maxwell, J. L., Keweshan, R. S. and Dunbar, M. W. (2011) Field-evolved resistance to Bt maize by western corn rootworm. *PLoS One* **6**: e22629.
- Gassmann, A. J., Shrestha, R. B., Jakka, S. R. K., Dunbar, M. W., Clifton, E. H., Paolino, A. R., *et al.* (2016) Evidence of resistance to Cry34/35Ab1 corn by western corn rootworm (Coleoptera: Chrysomelidae): root injury in the field and larval survival in plant-based bioassays. *Journal of Economic Entomology* **109**: 1872-1880.
- Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., *et al.* (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* **36**: 3420-3435.
- Heckel, D. G. (2012) Learning the ABCs of Bt: ABC transporters and insect resistance to *Bacillus thuringiensis* provide clues to a crucial step in toxin mode of action. *Pesticide Biochemistry and Physiology* **104**: 103-110.
- Huang, F. (2006) Detection and monitoring of insect resistance to transgenic Bt crops. *Insect Science* **13**: 73-84.
- Kim, D., Langmead, B. and Salzberg, S. L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**: 357.
- Kim, K. S. and Sappington, T. W. (2005) Genetic structuring of western corn rootworm (Coleoptera: Chrysomelidae) populations in the United States based on microsatellite loci analysis. *Environmental Entomology* **34**: 494-503.

- Kim, K. S., Stolz, U., Miller, N. J., Waits, E. R., Guillemaud, T., Sumerford, D. V., *et al.* (2008) A core set of microsatellite markers for western corn rootworm (Coleoptera: Chrysomelidae) population genetics studies. *Environmental Entomology* **37**: 293-300.
- Langmead, B. and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357-359.
- Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987-2993.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Ludwick, D. C. and Hibbard, B. E. (2016) Rootworm management: status of GM traits, insecticides and potential new tools. *CAB Reviews* **11**: 1-10.
- Ludwick, D. C., Meihls, L. N., Huynh, M. P., Pereira, A. E., French, B. W., Coudron, T. A., *et al.* (2018) A new artificial diet for western corn rootworm larvae is compatible with and detects resistance to all current Bt toxins. *Scientific Reports* **8**: 5379.
- Ludwick, D. C., Meihls, L. N., Ostlie, K. R., Potter, B. D., French, L. and Hibbard, B. E. (2017) Minnesota field population of western corn rootworm (Coleoptera: Chrysomelidae) shows incomplete resistance to Cry34Ab1/Cry35Ab1 and Cry3Bb1. *Journal of Applied Entomology* **141**: 28-40.
- McGaughey, W. H. (1985) Insect resistance to the biological insecticide *Bacillus thuringiensis*. *Science* **229**: 193-195.
- Meihls, L. N., Higdon, M. L., Ellersieck, M. and Hibbard, B. E. (2011) Selection for resistance to mCry3A-expressing transgenic corn in western corn rootworm. *Journal of Economic Entomology* **104**: 1045-1054.
- Meihls, L. N., Higdon, M. L., Ellersieck, M. R., Tabashnik, B. E. and Hibbard, B. E. (2012) Greenhouse-selected resistance to Cry3Bb1-producing corn in three western corn rootworm populations. *PLoS ONE* **7**: e51055.
- Meihls, L. N., Higdon, M. L., Siegfried, B. D., Miller, N. J., Sappington, T. W., Ellersieck, M. R., *et al.* (2008) Increased survival of western corn rootworm on transgenic corn within three generations of on-plant greenhouse selection. *Proceedings of the National Academy of Sciences* **105**: 19177-19182.
- Miller, N. J., Ciosi, M., Sappington, T. W., Ratcliffe, S. T., Spencer, J. L. and Guillemaud, T. (2007) Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation tolerance. *Journal of Applied Entomology* **131**: 378-385.

- Miller, N. J., Estoup, A., Toepfer, S., Bourguet, D., Lapchin, L., Derridj, S., *et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**: 992.
- Miller, N. J., Guillemaud, T., Giordano, R., Siegfried, B. D., Gray, M. E., Meinke, L. J., *et al.* (2009) Genes, gene flow and adaptation of *Diabrotica virgifera virgifera*. *Agricultural and Forest Entomology* **11**: 47-60.
- Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., *et al.* (2012) BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**: 1919-1920.
- Pardo-López, L., Soberón, M. and Bravo, A. (2013) *Bacillus thuringiensis* insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection. *FEMS Microbiology Reviews* **37**: 3-22.
- Rault, L. C., Siegfried, B. D., Gassmann, A. J., Wang, H., Brewer, G. J. and Miller, N. J. (2018) Investigation of Cry3Bb1 resistance and intoxication in western corn rootworm by RNA sequencing. *Journal of Applied Entomology* **142**: 921-936.
- Sonnhammer, E. L. L., Eddy, S. R., Birney, E., Bateman, A. and Durbin, R. (1998) Pfam: Multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Research* **26**: 320-322.
- Tay, W. T., Mahon, R. J., Heckel, D. G., Walsh, T. K., Downes, S., James, W. J., *et al.* (2015) Insect resistance to *Bacillus thuringiensis* toxin Cry2Ab is conferred by mutations in an ABC transporter subfamily A protein. *PLOS Genetics* **11**: e1005534.
- Vaughn, T., Cavato, T., Brar, G., Coombe, T., DeGooyer, T., Ford, S., *et al.* (2005) A method of controlling corn rootworm feeding using a protein expressed in transgenic maize. *Crop Science* **45**: 931-938.
- Walters, F. S., deFontes, C. M., Hart, H., Warren, G. W. and Chen, J. S. (2010) Lepidopteran-active variable-region sequence imparts coleopteran activity in eCry3.1Ab, an engineered *Bacillus thuringiensis* hybrid insecticidal protein. *Applied and Environmental Microbiology* **76**: 3082-3088.
- Wangila, D. S., Gassmann, A. J., Petzold-Maxwell, J. L., French, B. W. and Meinke, L. J. (2015) Susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to Bt corn events. *Journal of Economic Entomology* **108**: 742-751.
- Wu, T. D. and Watanabe, C. K. (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**: 1859-1875.
- Zhang, S., Cheng, H., Gao, Y., Wang, G., Liang, G. and Wu, K. (2009) Mutation of an aminopeptidase N gene is associated with *Helicoverpa armigera* resistance to *Bacillus thuringiensis* Cry1Ac toxin. *Insect Biochemistry and Molecular Biology* **39**: 421-429.

Zhao, Z., Meihls, L. N., Hibbard, B. E., Ji, T., Elsik, C. G. and Shelby, K. S. (2019) Differential gene expression in response to eCry3.1Ab ingestion in an unselected and eCry3.1Ab-selected western corn rootworm (*Diabrotica virgifera virgifera* LeConte) population. *Scientific Reports* **9**: 4896.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C. and Weir, B. S. (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**: 3326-3328.

Zukoff, S. N., Ostlie, K. R., Potter, B., Meihls, L. N., Zukoff, A. L., French, L., *et al.* (2016) Multiple assays indicate varying levels of cross resistance in Cry3Bb1-selected field populations of the western corn rootworm to mCry3A, eCry3.1Ab, and Cry34/35Ab1. *Journal of Economic Entomology* **109**: 1387-1398.

## VITA

Zixiao Zhao was born in the City of Quzhou, Zhejiang Province of China. In 2005 he enrolled College of Agriculture and Biotechnology at Zhejiang University (Hangzhou, China) and received bachelor's degree in applied biology with an emphasis in plant protection in 2009. In 2010 he joined the Department of Entomology at the University of Nebraska-Lincoln for a master's program. Advised by Dr. Nicholas J. Miller, he conducted research on the western corn rootworm (*Diabrotica virgifera virgifera* LeConte). Zixiao Zhao pursued a Ph.D. in plant, insect, and microbial science with a plant breeding, genetics and genomics emphasis area in the Division of Plant Sciences, University of Missouri in 2013. He joined the laboratories of Dr. Christine G. Elsik and Dr. Bruce E. Hibbard to continue research on the western corn rootworm. His research interests are insect genomics and transcriptomics, plant-insect interactions, mode of action of Bt (*Bacillus thuringiensis*) entomotoxins, and applying new sequencing and computational techniques in the areas of entomology. Zixiao Zhao was also appointed as adjunct instructor at Lincoln University (Jefferson City, Missouri) during his Ph.D. study where he taught a course of Integrated Pest Management to undergraduate students of the Department of Agriculture & Environmental Sciences.