A genetic and genomic study of resistance to 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor herbicides in Amaranthus tuberculatus (waterhemp)

Daniel Robert Kohlhase

Iowa State University
A genetic and genomic study of resistance to 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor herbicides in *Amaranthus tuberculatus* (waterhemp)

by

Daniel Robert Kohlhase

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Crop Production and Physiology (Weed Science)

Program of Study Committee:
Mike Owen, Major Professor
Michelle Graham
Jode Edwards
Allen Knapp

Iowa State University

Ames, Iowa

2016

Copyright © Daniel Robert Kohlhase, 2016. All rights reserved.
DEDICATION

To Dr. Mary Anne Bunda for encouraging and helping me begin these next stages of my academic life. Without your support and high expectations, my path would have been very different. Getting to know you and being guided by you showed me there are many ways of reaching a goal and beyond. Having you as a role model continues to challenge my way of thinking and motivates me to continuously better myself. Unfortunately, we didn’t have more time together. We love you and we miss you.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>1</td>
<td>GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Literature Review</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Thesis Organization</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>INHERITANCE OF 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (HPPD) INHIBITOR HERBICIDE RESISTANCE IN AMARANTHUS TUBERCULATUS (WATERHEMP)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Materials and Methods</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>AN RNA-SEQ DE NOVO TRANSCRIPTOME ASSEMBLY OF AMARANTHUS TUBERCULATUS (WATERHEMP) AND ANALYSES OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS RELATED TO 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (HPPD) INHIBITOR HERBICIDE RESISTANCE</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Materials and Methods</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Acknowledgements</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>GENERAL CONCLUSION</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>90</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Mike Owen, and my committee members Dr. Michelle Graham, Dr. Jode Edwards, and Dr. Allen Knapp for their guidance, support, and patience. I deeply grateful for all of the assistance I have received that made these projects possible and enjoyable, especially Dr. Mike Owen’s lab, Dr. Michelle Graham’s lab, and Dr. Jamie O’Rourke’s lab. In addition, I would like to thank the Department of Agronomy for making my master’s degree experience at Iowa State University so enjoyable. Finally, I would like to thank my family and friends. Words cannot fully express my appreciation for the constant love, support, and encouragement I have received over the course of this project.
ABSTRACT

Waterhemp (*Amaranthus tuberculatus*) is an agronomic weed found throughout the Midwestern United States. Without proper management waterhemp has the potential to cause yield losses up to 74% and 56% in maize (*Zea mays* L.) and soybean (*Glycine max* (L.) Merr.), respectively. There are various characteristics of waterhemp that contribute to increasing prevalence of waterhemp in agriculture including the ability to evolve herbicide resistance. There are six herbicide sites of action that have been evolved in waterhemp. In 2011, waterhemp was discovered to be resistant to p-hydroxyphenylpyruvate-dioxygenase (HPPD, EC 1.13.11.27) inhibitor herbicides. The objectives of my research were to 1) identify the number of genes and describe the allelic expression involved with HPPD-herbicide resistance in waterhemp and 2) examine the transcriptomic expression response of HPPD-herbicide resistance in waterhemp.

To examine the inheritance of the resistance trait, we reciprocally crossed a known HPPD resistant waterhemp population with a known HPPD susceptible waterhemp population and established a pseudo-F₂ generation. We challenged the parent, F₁ and pseudo F₂ generations against four rates of an HPPD-inhibiting herbicide. The HPPD-resistance trait was determined to be polygenic. Additionally, these data suggest the number of involved resistance genes was found to increase at higher rates of the herbicide with at least one dominant allele at each major locus. This study has confirmed previous reports describing the inheritance of HPPD resistance to be complex while introducing new descriptions of the response of HPPD resistant waterhemp to HPPD herbicides.
To examine the transcriptomic response of resistant waterhemp, we treated and mock-treated HPPD-herbicide resistant and susceptible waterhemp populations with mesotrione and collected leaf samples at three, six, twelve, and twenty-four hours after treatment (HAT). We performed a *de novo* transcriptome assembly using all sample sequences. This de novo assembly was then used to measure gene expression differences between genotypes, treatments and time points. These data suggest a rapid response of HPPD-herbicide resistant and susceptible waterhemp genotypes to the HPPD-inhibiting herbicide mesotrione. Furthermore, little overlap was found among the differentially expressed transcripts expressed by each genotype. We also identified the possibility of overlapping gene networks in response to other herbicides. We have made available the raw sequences, and assembled sequences with complete annotations for continued use by the weed science community.
CHAPTER 1: GENERAL INTRODUCTION

Literature Review

Introduction

It is more important to have a full understanding of herbicide resistances due to the challenge of herbicide-resistant weeds increasing prevalence in global agriculture. *Amaranthus tuberculatus* (Moq.) J.D. Sauer (waterhemp) represents a model weed to study herbicide resistance given that resistance to six herbicide sites of action have been documented (Heap 2016). Waterhemp has become an increasing problem for crop production in the Midwest over the past 20-30 years with the increased adoption of herbicides (Powles et al. 1997, Steckel 2007). In 2009, 4-hydroxyphenylpyruvate dioxygenase (HPPD)-inhibitor herbicides were added to the waterhemp herbicide resistance list and now resistant populations have been reported in several Midwestern states (Heap 2016). The first cases of HPPD herbicide resistance resulted from an average of 6 applications of an HPPD-inhibiting herbicide before fields were identified with a putative resistant waterhemp population. Previous research has described the waterhemp resistance to HPPD-inhibitor herbicides attributable to increased cytochrome P450 monooxygenase activity resulting in a high level of HPPD herbicide metabolism (Ma et al. 2013). Due to the non-target-site mechanism of HPPD resistance, specifically metabolic-based, this trait is likely under polygenic control (Mithila and Godar 2013). Huffman et al. (2015) described the inheritance of HPPD resistance in a waterhemp population from Illinois to be complex although there are no reports specifically describing the gene expression of HPPD resistance in waterhemp.
Biology

There are many species of *Amaranthus* native to North America. The two native *Amaranthus* spp. of the Midwest are *Amaranthus rudis* J.D. Sauer (common waterhemp) and *Amaranthus tuberculatus* (Moq.) J.D. Sauer (tall waterhemp). Pratt and Clark (2001) suggested a single polymorphic species of waterhemp and referred to it as *A. tuberculatus*. Waterhemp is a summer annual that has the potential to grow up to 3.66 meters but typically grows to 1.22 -1.52 meters in a field (Steckel 2007). Identifiable characteristics include green (sometimes red) glabrous stems and lanceolate leaves that appear waxy or glossy. There are many favorable biological traits that contribute to the effectiveness of waterhemp as an agronomically important weed including seed germination and emergence patterns, plant growth rate, the dioecious reproductive habit, and the amount of offspring produced by a single female plant. Extended germination allows for several opportunities of establishment throughout the growing season (Hartzler et al. 1999). High growth rates, potentially up to 2.45 cm per day, help the plants successfully compete for necessary resources. Waterhemp is dioecious; the male and female flowers are on separate plants (Sauer 1955). A dioecious flowering system forces the species to cross-pollinate for every fertilization event. Cross-pollination is advantageous because this stimulates genetic diversification and the likely sharing of desirable genetic traits. Waterhemp has the potential to produce more than 1 million seeds on a single female plant in the field. The abundant seed production contributes to the probability of adaptation when exposed to specific selection pressures such as that imposed by herbicides (Trucco et al. 2005).
**Herbicide Resistance**

Herbicide resistance, as defined by the Weed Science Society of America (WSSA), is “the inherited ability of a plant to survive and reproduce following exposure to a dose of herbicide normally lethal to the wild type” (Vencill et al. 2012). The transition of a herbicide-susceptible population to a primarily herbicide-resistant population reflects the contribution of a number of factors such as species genetics, biology, and management (Powles and Yu 2010). Herbicide resistance is inherited in future generations through the transmission of genetic material. Biological factors include characteristics such as dioecious reproduction resulting in outcrossing, opportunistic germination habit and high seed production which can all contribute to an increased probability that population will evolve herbicide resistance. Proper weed management involves a number of considerations including the herbicide(s) used, frequency of use, and rate of application. An important driving force of herbicide resistance evolution is the selection pressure to the weed population resulting from herbicide programs used in modern farm practices (Owen and Zelaya 2005). It is important to understand the herbicides being used in a weed management plan because a lack of diversity in herbicide mode of action (MOA) can result in a more rapid transition to a herbicide-resistant population (Sutton et al. 2002).

The two primary classes of herbicide resistance are target-site and non-target-site resistance. Target-site resistance (TSR) can be a change in the enzyme binding site for the herbicide causing the herbicide to become ineffective. A mutation of the herbicide target site changes the 3D structure of the enzyme resulting in the inability of the herbicide to properly bind (Délye et al. 2013). TSR could also be an overexpression of
the target enzyme (Powles and Yu 2010). This resistance occurs when the regular dose of herbicide does not sufficiently inhibit enough of the target enzyme to obstruct an essential physiological process that the plant needs to survive. Non-target-site resistance (NTSR) reduces the amount of herbicide that successfully reaches the intended target-site. One mode of NTSR is reduced herbicide penetration into the plant. Another mechanism is reduced translocation of the herbicide to the target enzyme. Also enzymes such as cytochrome P450s (P450), glutathione-S-transferases, glycosyltransferases, esterases, and hydrolases can increase metabolism of the herbicide (Délye 2013).

Cytochrome P450s encompass a large and diverse enzyme superfamily (Hannemann et al. 2007). There are many genes that code for multiple types of P450s. The number of genes associated with the coding of P450 differs by species. Two hundred seventy two genes are associated with P450s in Arabidopsis thaliana (L.) Heynh and there are 458 in Oryza sativa L. Cytochrome P450s differ in reactive sites and can be organized into various families. Plant species cumulatively contain 59 cytochrome P450 families (Schuler and Werck-Reichhart 2003). These heme-dependent enzymes activate molecular oxygen using electrons transferred from NADPH catalyzing a mono-oxygenation reaction with the herbicide. The reaction reduces the phytotoxicity of the herbicide allowing the plant to further metabolize the chemical (Délye 2013, Werck-Reichhart et al. 2000). Hydroxylations are one of the most common reactions among the large array of reactions that are catalyzed by cytochrome P450s and is thought to be the reason for the effective metabolism of mesotrione, an HPPD (p-hydroxyphenylpyruvate dioxygenase (HPPD, EC 1.13.11.27) inhibiting herbicide, in Zea mays L. (corn) (Cobb and Kirkwood 2000, Ma et al. 2013). There is supporting evidence that a herbicide-
resistant waterhemp population from Illinois uses similar mechanisms for mesotrione detoxification (Ma et al. 2013).

**HPPD Herbicides**

Herbicides that inhibit HPPD were discovered in the early 1980’s and first commercialized in 1993 (Mitchell et al. 2001). Herbicides that target HPPD are comprised of several families including isoxazoles, triketones, and pyroxazoles. Reduction of p-hydroxyphenylpyruvate (HPP) to homogentisate (HGA) is catalyzed by HPPD. HGA is a precursor in the biosynthesis of prenylquinones, which include vital compounds such as vitamin E and plastoquinones (Matringe et al. 2005). Plastoquinone is an electron acceptor in photosystem II and is involved in carotenoid biosynthesis. Reduced synthesis of plastoquinone and carotenoids inhibits proper photon processing. This results in an accumulation of free radicals resulting in photooxidation and eventually pigment damage in the form of a bleached meristems occurs, which has become an identifying symptom of HPPD-inhibitor herbicides (Hamprecht and Witschel 2012, Ma et al. 2013).

HPPD-inhibiting herbicides represent the latest herbicide group with a new mechanism of action to be commercially available. These herbicides are represented by products containing active ingredients such as isoxaflutole, mesotrione, tembotrione, and topramezone (Hausman et al. 2011, McMullan and Green 2011). Mesotrione is widely used in corn production because of the reduction in crop injury due to cytochrome P450-based differential metabolism (Hamprecht and Witschel 2012). The recommended application rate ranges from 100 to 225 g ha\(^{-1}\) as a pre-emergence and 70 to 150 g ha\(^{-1}\) as a post-emergence herbicide (Matringe et al. 2005).
Resistance to HPPD-inhibiting herbicides was initially discovered in 2009 when seed corn growers in McLean County, Illinois, USA found waterhemp that was unresponsive to applications of tembotrione and mesotrione (Hausman et al. 2011). The same year resistance was also confirmed in a population of waterhemp that did not respond to mesotrione and atrazine applications in a seed corn field located in Henry Country, Iowa, USA (McMullan and Green 2011). Two years later HPPD resistant waterhemp was discovered in Nebraska (Heap 2016). Preliminary studies suggested the waterhemp to have non-target-site resistance attributable to increased metabolic activity of the cytochrome P450 monooxygenase (Ma et al. 2013). The details of the metabolic resistance (broad specificity of a single cytochrome P450 or low specificity of multiple cytochrome P450s) in waterhemp has yet to be determined. The only other weed with documented resistance to HPPD-inhibiting herbicides is *Amaranthus palmeri* S. Watson (Palmer amaranth), confirmed in Kansas and Nebraska (Heap 2016).

**Genetics – Inheritance and Sequencing**

An understanding of the mode of inheritance is helpful when trying to manage the spread of herbicide resistance throughout a susceptible weed population (Okada and Jasieniuk 2014). Important factors that are considered when examining resistance on a genetic level include the number of genes controlling the resistance trait and the dominance of the trait. When a trait is controlled by a single gene it is considered to be monogenic (Mithila and Godar 2013, Sleper and Poehlman 2006). A single trait that is coded by multiple genes is considered to be polygenic. Genes are referred to as dominant when they are expressed regardless of zygosity. Recessive genes are characterized by the omitted expression in the presence of a dominant allele. An additional important genetic
factor is whether the trait is maternally inherited. Nuclear inherited genes are transmitted through gametes of either sex while maternally inherited genes are transferred only through female plants. There are multiple mechanisms of maternal inheritance (Roach and Wulff 1987). One mechanism typically associated with maternal inheritance is the transmission of traits found on the genome of the chloroplast or mitochondria, which is described as cytoplasmic inheritance.

Until recently, scientists were unable to study non-model organisms at a genetic level due to financial and technological barriers. Next generation sequencing (NGS) is opening the possibility of a range of non-model organisms to be studied at a genomic level at a practical cost (Ekblom and Galindo 2011). Lee et al. (2009) sampled the waterhemp genome by isolating DNA and using Roche 454 pyrosequencing (Roche Sequencing, Pleasanton, CA, USA) after noticing an absence of genomic resources in the field of weed science (Lee et al. 2009). Shortly after, Riggins et al. (2010) added to the Amaranthus spp. genomic resource, in particular waterhemp, by sequencing the transcriptome using Roche GS-FLX 454 pyrosequencing (Roche Sequencing, Pleasanton, CA, USA) and characterizing herbicide target-site genes identified from other species (Riggins et al. 2010). The transcriptome, or all of the expressed transcripts in a cell, is important because it represents the functional activity in the genome. NGS can be useful when trying locate genes that control advantageous plant characteristics by sequencing the transcriptome when the characteristics of interest are expressed (Strickler et al. 2012).

The next advancement in the study of transcriptomes and gene expression is RNA sequencing (RNA-seq). In addition to which genes are expressed, RNA-seq is able to quantify the level of gene expression (Wang et al. 2009). Solexa Illumina (Illumina Inc.,

---

Wang et al. 2009

Lee et al. 2009

Ekblom and Galindo 2011

Roach and Wulff 1987

Strickler et al. 2012
San Diego, CA, USA) is one of the leading platforms used in genetic studies of non-model organisms. The platform begins the assembly process by converting mRNA to fragments of cDNA while adding adapters to either one or both ends of the fragment. The cDNA fragments are then sequenced from one or both ends of the fragment. Paired-end (PE) sequencing will typically cover 75-150 base pairs (bp) from both ends of the cDNA in a “forward-reverse” orientation. The duplicated sequences are aligned thus helping increase the accuracy and length of the reads. Non-model organisms typically do not have an available reference genome to assist in assembling the sequenced data, limiting the available options to a de novo assembly (Martin and Wang 2011, Strickler et al. 2012, Wang et al. 2009). There are multiple assembly methods available that will produce de novo assemblies from RNA-seq data including Trans-ABySS (Robertson et al. 2010), Velvet-Oases (Schulz et al. 2012), and SOAPdenovo-trans (Xie et al. 2014) (Haas et al. 2013). Trinity (Grabherr et al. 2011) is another method that has proven be reliable and highly effective. Trinity divides the data, generated by RNA-seq, to construct de Bruijn graphs from which transcripts and alternatively spliced isoforms can be reconstructed. After each assembly, quality analyses need to be performed to confirm the validity of the transcriptome. Once a final assembly is accepted, analyses can be performed such as differential expression, characterization, and annotation of the transcriptome.

**Summary**

There are many traits that contribute to the success of waterhemp as a pest in crop production as well as the ability of waterhemp to evolve resistance to herbicides. Prolific seed production, high growth rates, and dioecious habit of waterhemp have all contributed to the evolution of resistance towards six sites-of-action (Hartzler et al. 1999,
Heap 2016, Horak and Loughin 2000, Steckel et al. 2003). Herbicide resistance can be broadly separated into two categories of mechanisms depending on the ability of the herbicide to reach the intended target site at an effective dose, target-site (TSR) and non-target-site resistance (NTSR) (Powles and Yu 2010). Enzymes, such as cytochrome P450s, glutathione-S-transferases, glycosyltransferases, esterases, and hydrolases, metabolize herbicides and are an important mechanisms of NTSR (Délye 2013).

In 2011, two distinct populations of waterhemp were discovered to have evolved mesotrione resistance (Hausman et al. 2011, McMullan and Green 2011). Cytochrome P450 has been cited as the mechanism of resistance in mesotrione herbicide resistant waterhemp (Ma et al. 2013). Mesotrione belongs to the p-hydroxyphenylpyruvate dioxygenase (HPPD) inhibiting herbicides (Mitchell et al. 2001). The inhibition of HPPD will prevent the synthesis of plastoquinones and carotenoids. The inability of the plant to process photons will cause free radical damage and the bleaching of meristematic tissue.

Favorable plant characteristics can be studied from a genetic or genomic approach for a better understanding of the characteristic. The genetic approach helps to identify how the trait is inherited and how it moves between generations. The genomic approach examines the trait at a molecular level and helps to identify specific genes and gene expression responsible for the trait. The development of RNA-seq provides the opportunity to study non-model organisms that may not have a large amount of genomic resources, such as a reference genome (Ekblom and Galindo 2011). We will examine the mesotrione resistance trait in waterhemp from both a genetic and genomic approach. The genetic study will contribute to our understanding of the spread and development of this resistance while the genomic study will help increase our understanding of the
waterhemp molecular stress response to herbicide and potentially identify important genes that play a major role in herbicide resistance.

**Thesis Organization**

This thesis is organized as two manuscripts suitable for the submission to scientific journals. There is a general introduction and general conclusion located before and after the two manuscripts, respectively. The first manuscript entitled, “Inheritance of 4-hydroxyphenylpyruvate Dioxygenase Inhibitor Herbicide Resistance in *Amaranthus Tuberculatus*” is suitable for submission to *Plant Science*. All individuals listed as authors have made contributions to this section: experimental design (D.K., M.O. and J.E.), data collection (D.K.), data analysis (D.K. and J.E.), and manuscript preparation (D.K. and M.O.). The second manuscript entitled, “An Rna-Seq De Novo Transcriptome Assembly of *Amaranthus Tuberculatus* and Analyses of Differentially Expressed Transcripts Related to 4-Hydroxyphenylpyruvate Dioxygenase Inhibitor Herbicide Resistance” is suitable for submission to *BMC Genomics*. All individuals listed as authors have made contributions to this section: experimental design (D.K., M.O. and M.G.), data collection (D.K. and M.G.), data analysis (D.K. and M.G.), and manuscript preparation (D.K., M.O. and M.G.). Both manuscripts include an abstract, introduction, materials and methods, results, discussion, conclusion, references, tables, and figures. The section order within each manuscript is journal specific.
References


CHAPTER 2: INHERITANCE OF 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE INHIBITOR HERBICIDE RESISTANCE IN *AMARANTHUS TUBERCOULATUS*

A paper to be submitted to *Plant Science*

Daniel Kohlhase\(^a\), Jode Edwards\(^b\), Micheal Owen\(^a\)

**Abstract**

Waterhemp (*Amaranthus tuberculatus*) is a weed prevalent in the Midwestern United States and can cause yield losses up to 74% in maize (*Zea mays* L.) and 56% in soybean (*Glycine max* (L.) Merr.). An important adaptive trait found in waterhemp is the ability to evolve herbicide resistance and waterhemp populations have evolved resistance to six herbicide sites of action. In 2011, two waterhemp populations were discovered resistant to p-hydroxyphenylpyruvate-dioxygenase (HPPD, EC 1.13.11.27) inhibitor herbicides. Quickly after the discovery of HPPD-herbicide resistance, studies were performed to determine the mechanism of resistance and describe the inheritance of the herbicide resistance trait. We reciprocally crossed a known HPPD resistant waterhemp population with a known HPPD susceptible waterhemp population and established a pseudo-F\(_2\) generation. We challenged the parent, F\(_1\) and pseudo F\(_2\) generations against four rates of an HPPD-inhibiting herbicide (mesotrione). Our analyses focused on describing the number of, and characterizing the allelic expression of the genes responsible for the HPPD-resistance in waterhemp. Our results suggest the HPPD-resistance trait to be polygenic. Furthermore, the number of genes involved with the herbicide resistance increase at higher rates of the herbicide. These data indicated at least one dominant allele is needed at each major locus in order to confer resistance. By using
a different waterhemp population and methodology, this study confirmed the reported complex inheritance of HPPD resistance while suggesting new details in the response of HPPD resistant waterhemp to HPPD herbicides.

1. Introduction

Waterhemp (*Amaranthus tuberculatus* (Moq.) J.D. Sauer) is a major weed problem in maize (*Zea mays* L.) and soybean (*Glycine max* (L.) Merr.) production systems across the Midwest United States. This highly competitive summer annual weed has the potential to reduce corn and soybean yields up to 74% and 56%, respectively [1]. Even though waterhemp is native to the Midwest, this species has increased in prominence due to changes in herbicide use and tillage practices [2]. Waterhemp is a dioecious diploid (2n = 32); the female and male reproductive structures are found on separate plants [3]. The dioecious nature of waterhemp makes this species an obligate outcrosser. Mandatory outcrossing for every fertilization event can be advantageous for this species because the consistent genetic recombination provides ample opportunities for the movement and development of valuable traits, such as herbicide resistance, among and between waterhemp populations. Additional biological traits that are associated with the competitive success of waterhemp include prolific seed production [4], extended germination period [5], and rapid growth rate [6].

Since the commercialization of chemical weed control in the late 1940’s, herbicides have made a large impact on the approach farmers take to weed management practices [7]. Shortly after the introduction of herbicides into the market, the first case of evolved herbicide resistance was documented [8]. Currently there are 470 documented cases (case is defined as an individual species by specific site of action) of herbicide
resistance around the world and the number continues to grow [9]. Herbicide resistance is classified by two major categories of mechanisms, target-site resistance (TSR) and non-target-site resistance (NTSR) [10]. In TSR, the herbicide reaches the intended target site but is ineffective because of an inability for the chemical to properly bind or because of overexpression of the target enzyme; for NTSR, the herbicide may be unable to reach the intended target site because of decreased penetration, inhibited translocation, sequestration, or metabolism. Due to the specificity of herbicide design, TSRs are typically the result of a single mutation and accordingly monogenic (controlled by a single gene) and dominant or semi-dominant [11]. NTSRs are typically polygenic (controlled by multiple genes); several studies have demonstrated metabolic herbicide resistance to be polygenic [12–15].

In the early 1990’s 4-hydroxyphenylpyruvate dioxygenase (HPPD, EC 1.13.11.27) inhibitor herbicides were commercialized [16]. HPPD is responsible for the conversion of 4-hydroxyphenylpyruvate (4-HPP) to homogentisate (2,5-dihydroxyphenylacetate; HGA), which is the aromatic precursor of plastoquinones. Due to the connection between plastoquinones and carotenoid biosynthesis, the herbicidal activity of HPPD is typically characterized by bleaching of new tissue. Mesotrione (Herbicide Group 27) is a selective HPPD-inhibitor herbicide widely used for broadleaf control in corn [17]. Almost 20 years after commercialization, the first cases of resistance to HPPD inhibitors were simultaneously verified in Iowa and Illinois when waterhemp populations did not respond to regular herbicide applications in seed corn production fields [18,19]. The only other species to have confirmed resistance to HPPD inhibitors is Palmer amaranth (Amaranthus palmeri S. Watson) [9]. The mechanism of
resistance in the Illinois waterhemp population was theorized to be the result of increased mesotrione metabolism by cytochrome P450 monooxygenase (CYP450, EC 1.14.14.1) and is an example of NTSR [20]. The same Illinois waterhemp population was used to investigate the inheritance of the resistance and the inheritance was reported to be “complex” [21].

Even though the inheritance of HPPD resistance was examined in the Illinois waterhemp population, the same resistance mechanism cannot be assumed for other resistant waterhemp populations. For example, two Illinois atrazine (Herbicide Group 5) resistant waterhemp populations were distinguished as having two distinct resistance mechanisms [22]. In this experiment, we cross a known susceptible waterhemp population and a confirmed HPPD-resistant waterhemp population as the parent generation to characterize the mode of inheritance of HPPD herbicide resistance through two subsequent generations.

2. Materials and Methods

2.1 Parent Population Collection

The mesotrione-susceptible waterhemp population (designated as CFS for Curtiss Farm mesotrione-susceptible) was collected from the Curtiss Farm at Iowa State University (Ames, IA, USA) in 2006. The mesotrione-resistant waterhemp population (designated as HCR for Henry County mesotrione-resistant) was collected from a seed corn production field in Henry County, IA, USA in 2009 following reports of waterhemp surviving mesotrione applications [18]. The population was characterized as resistant to three herbicide modes of action; HPPD inhibitors (Herbicide Group (HG) 27),
photosystem-II inhibitors (HG 5), and acetolactate synthase (ALS) inhibitors (HG 2) [18].

2.2 Resistant Parent Recurrent Selection

HCR was cycled through three cycles of recurrent selections (HCR₁, HCR₂, HCR₃) by mesotrione with a rate of 421.44 g ai ha⁻¹ used in the final cycle, which represents 4x the suggested field rate (Callisto®, Syngenta Crop Protection, Inc., Greensboro, NC 27419-8300). Seeds were stratified to help break seed dormancy. Petri dishes containing the seed sample and a very thin layer of water were placed in cold storage at 6°C for two weeks. The lids of the petri dishes were removed and the petri dishes were transferred for two days to a dryer set at 45°C. The seeds were then planted in 0.31m x 0.31m flats using a 4:1 mixture of Sunshine Mix #1/LC1 potting soil (Sun Gro Horticulture, Agawam, MA, USA) to sand ratio. Plants were grown in a greenhouse set to 24°C and were watered as needed; sunlight was supplemented with 600-1,000 µmol m⁻² s⁻¹ photosynthetic photon flux density (PPFD) of artificial light set to a 14 hour photoperiod. Peters® Excel fertilizer (15-5-15; ICL Specialty Fertilizers, Geldermalsen, Netherlands) was used as needed by injecting the water soluble fertilizer into the water supply when the plants were watered. When the plants reached an average height of 7.6 cm, they were treated with mesotrione (105.36 g ai ha⁻¹), 1% (v/v) crop oil concentrate (COC), 2.5% (v/v) urea ammonium nitrate solution (UAN), and water using 191.76 L ha⁻¹ carrier volume through a CO₂-powered spray chamber equipped with TeeJet® 80015EVS nozzles (Spraying Systems Co., Wheaton, IL, USA). The surviving plants were grown out through reproduction and the seeds were collected. The following year the process was repeated with the same mesotrione treatment using the seed collected
from the surviving plants. The third year the process was repeated but plants were treated with 421.44 g ai ha\(^{-1}\) of mesotrione and adjuvants.

2.3 Generation of F\(_1\) and F\(_2\) families

2.3.1 Production of F\(_1\) families

The seeds from the CFS and HCR\(_3\) waterhemp populations were stratified and germinated in flats in the greenhouse as described in section 2.2. The HCR\(_3\) plants were treated with mesotrione (105.36 g ai ha\(^{-1}\)) as described in section 2.2 when the plants reached an average height of 7.6 cm to ensure a homogeneous group of HPPD-resistant waterhemp was used for the cross. The surviving plants were transplanted into peat pots two weeks after treatment. Waterhemp is a short-day plant, which means the plants will initiate flowering when exposed to shorter photoperiods regardless of plant age and size [23,24]. In order to determine the sex, plants were induced to flower by placing them under a black plastic cover to control the exposure to light. The daily light exposure was reduced to 11-12 hours for approximately 7 to 12 days until plants initiated flowers. The populations were then segregated by sex and then exposed to the normal day length photoperiod so they could revert to vegetative growth.

After sex segregation in the greenhouse, the plants were transplanted to the field in the first week of June 2013. The field was a Clarion loam (21% clay, 34% silt, 45% sand) along the Bemis moraine with a two to six percent slope, 6.2 pH, and 3.5% organic matter (http://websoilsurvey.nrcs.usda.gov/). Two large (2.1m x 6.1m x 1.8m) polyethylene mesh pollination tents (Redwood Empire Awning, Santa Rosa, CA, USA) were erected in opposite corners of the field. Landscape fabric covered the soil inside the tents to suppress volunteer weeds. Each tent contained 200 plants, 100 of each sex. The
tent was divided into four quadrants with each quadrant containing 50 plants of one sex, resulting in two male quadrants and two female quadrants (Figure 1). The plants in each quadrant were arranged in a 5 x 10 plant grid to maximize space between plants. The quadrants were alternated by sex in order to achieve maximum pollen distribution to the female plants. The first tent had mesotrione-resistant females paired with mesotrione-susceptible males (HCR₃-♀ x CFS-♂), hence forth designated HCR₃-♀, and the second tent contained herbicide-susceptible females and herbicide-resistant males (CFS-♀ x HCR₃-♂), hence forth designated HCR₃-♂. The transplants were hand-watered until new growth was observed indicating successful establishment in the field; plants were not hand-watered afterwards as rainfall provided sufficient water to support growth. The area surrounding both tents was hand-weeded throughout the growing season to eliminate volunteer weeds and reduce the possibility of pollen contamination.

After the seeds representing the first filial generation (F₁) were physiologically mature, individual female parent plants were excised at ground level, placed in paper bags, and air dried for 2-3 weeks. Seeds from an individual plant represented an F₁ family. Initially the plants were hand-threshed and seeds were separated from larger vegetative material using a sieve. Threshed seeds were further cleaned using a South Dakota Seed Blower (Seedburo Equipment Company, Chicago, IL, USA). The amount of cleaned seed for each F₁ family was estimated to determine which families produced enough seed for future research including the development of the F₂ generation as well as subsequent mesotrione evaluations. A total of 98 F₁ families (64 HCR₃-♀ and 34 HCR₃-♂) were established.
2.3.2 Field Production of the F\textsubscript{2} families

In June 2014, 97 F\textsubscript{1} families were selected, stratified as described in section 2.2, and germinated in individual 10.16 cm round plastic pots using Sunshine Mix #1/LC1 (Sun Gro Horticulture, Agawam, MA, USA) potting soil. Plants were grown in a greenhouse set to 24°C and were watered as needed; sunlight was supplemented with 600-1,000 µmol m\textsuperscript{-2} s\textsuperscript{-1} PPFD of artificial light set to a 14 hour photoperiod. 16 seedlings per F\textsubscript{1} family were transplanted into Ray Leach Cone-tainers\textsuperscript{TM} (Stuewe & Sons, Inc., Tangent, Oregon, USA) when they reached an average height of 2.5 cm approximately 2-3 weeks after transplanting, the plants were induced to flower using the black plastic cover as described in section 2.3.1 allowing plants to be segregated by sex. Waterhemp is dioecious which prohibits self-pollination so we established a pseudo-F\textsubscript{2} generation (hence forth called F\textsubscript{2}) by inter-mating the F\textsubscript{1} families, a procedure modified from Huffman et al. (2015). Each inter-mated F\textsubscript{1} family was replicated three times resulting in a total of 291 “self-crosses”. Each replicate was then treated as an individual from an F\textsubscript{2} family. Three F\textsubscript{1} plants (two female and one male) were transplanted into a 25.4 cm round pot and the pots were immediately transplanted into the field using a completely randomized experiment design. Transplanting occurred the third week of July 2014. Rows of plastic mulch spanning the width of the field were established the week before transplanting and spaced 3.05m apart on center to maximize distance between rows to reduce the potential for pollen contamination; holes were established in the plastic for the 25.4 cm round pots which were spaced 4.26m apart on center.
Each transplanted pot was covered just prior to anthesis with a custom made 1.22m x 1.83m Delnet® Pollination™ bag (DelStar Technologies, Inc., Middletown, DE, USA) in order to eliminate pollen contamination between F\textsubscript{1} families. The Delnet bags were supported by 4, 1.27cm x 1.52m PVC pipes with caps on the exposed ends, hammered in the ground in a 0.46m x 0.46m square surrounding each pot. At the end of the growing season, seeds representing the second filial generation (F\textsubscript{2}) were harvested and cleaned in the same manner as the F\textsubscript{1} seeds. When harvesting, seeds derived from the same pot were bulked. A total of 274 F\textsubscript{2} families (179 HCR\textsubscript{3}-♀ and 95 HCR\textsubscript{3}-♂) were collected.

2.4 Parent Generation Mesotrione Dose Response

Seeds from the two parent populations (CFS and HCR\textsubscript{3}) were planted in 30 individual 11 cm round pots using a 4:1 mixture of Sunshine Mix #1/LC1 potting soil (Sun Gro Horticulture, Agawam, MA, USA) to sand ratio and 0.25 tsp of Osmocote Flower Food Granules (14-14-14) (The Scotts Miracle-Gro Company, Marysville, OH, USA). Plants were grown in a greenhouse set to 24°C and were watered as needed; sunlight was supplemented with 600-1,000 µmol m\textsuperscript{-2} s\textsuperscript{-1} PPFD of artificial light set to a 14 hour photoperiod. The seedlings were thinned to 2 plants per pot and placed in greenhouse in a split-plot, randomized complete block design.

Plants were treated with mesotrione as described in section 2.2 when plants reached an average height of 7.6 cm. Treatments included 5 rates of mesotrione, based on a 1x field rate of 105.36 g ai ha\textsuperscript{-1} included 0.5x (52.68 g ai ha\textsuperscript{-1}), 1x, 2x (210.72 g ai ha\textsuperscript{-1}), 4x (421.44 g ai ha\textsuperscript{-1}) and 8x (842.88 g ai ha\textsuperscript{-1}) plus a control treatment that was sprayed with water. Each treatment was replicated 5 times. Treatments were evaluated for tissue
damage 7, 14, and 21 days after treatment (DAT) using a scale based on visual injury percent where healthy tissue represents 0% injury with increasing visual injury percent for bleached, necrotic tissue to plant death which represents 100% injury.

2.5 Parent, F₁, and F₂ Herbicide Injury Evaluations

Cumulatively, 112 families were screened for herbicide injury; 46 families were from the F₁ generation and 66 families were from the F₂ generation. While the F₂ seeds were being produced in the field, we screened 24 families from the F₁ generation without the F₂ replicates derived from those F₁ families. Within the F₁ generation, 25 families were from the HCR₃-♀ parent cross and 21 families were from the HCR₃-♂ parent cross. Within the F₂ generation, 36 families were from the HCR₃-♀ parent cross and 30 families were from the HCR₃-♂ parent cross. Families were selected to be screened if a F₁ family and all three F₂ replicates derived from that F₁ family had enough seed to complete the experiment. A maximum of 20 families were evaluated for herbicide injury at the same time in the greenhouse due to logistics.

Seeds were stratified as described in section 2.2 and then planted in 10.16 cm round plastic pots using a 4:1 mixture of Sunshine Mix #1/LC1 potting soil (Sun Gro Horticulture, Agawam, MA, USA) to sand ratio and 0.25 tsp of Osmocote Flower Food Granules (14-14-14) (The Scotts Miracle-Gro Company, Marysville, OH, USA). Twenty-five pots were prepared for each family. Plants were grown in a greenhouse set to 24°C and were watered as needed; sunlight was supplemented with 600-1,000 µmol m⁻² s⁻¹ PPFD of artificial light set to a 14 hour photoperiod. The pots were placed in the greenhouse in a split-plot, randomized complete block design as described in section 2.4.
Plants were treated with mesotrione as described in section 2.2 at an average height of 7.6 cm. Treatments included water, mesotrione at a 0.5x field rate (52.68 g ai ha$^{-1}$), 1x (105.36 g ai ha$^{-1}$), 2x (210.72 g ai ha$^{-1}$), and 4x (421.44 g ai ha$^{-1}$). Each treatment was replicated 5 times. Plants were evaluated 7, 14, and 21 DAT for herbicide response using the same scale of visual injury percent described in section 2.4.

2.6 Data Analysis

All statistical analyses were performed using Statistical Analysis Software, SAS 9.3 (SAS Institute Inc., NC, USA). All data for the three plant generations (parent, F$_1$, and F$_2$) were analyzed separately at each DAT using beta regression to test for significant differences among families within each generation (PROC GLIMMIX).

2.6.1 Parent, F$_1$, and F$_2$ Cluster Analyses

There are no standards of herbicide injury response that clearly define an injury response threshold thus allowing plants to be classified as susceptible or resistant to a herbicide. Therefore, cluster analyses was performed in order to establish resistance thresholds to classify plants as susceptible or resistant within the mesotrione rates used in this experiment. The cluster analysis tests for distinct categories of herbicide injury responses, which can be indicative of the number of genes involved in the resistance trait. Also, creating a resistance threshold for what is considered susceptible or resistant is necessary in order to perform a chi-square ($\chi^2$) goodness-of-fit (GOF) test. The data were tested for natural clustering and optimum cluster numbers (PROC CLUSTER) based on damage scores within a plant generation, herbicide rate, and DAT combination. This analysis established the number of clusters the data naturally segregates, frequency of the
cluster segregation ratios, and the mean damage score within those clusters. The frequency of the cluster segregation ratios was used as the frequency of resistant to susceptible (R:S). The cluster means were used in determining the resistance thresholds for this experiment. No patterns of natural clustering within the data were observed so the cluster analysis was restricted to a maximum number of clusters, forcing the analysis to find cluster means within the given maximum number of clusters and then dividing the rest of the data based on distance from those means. The cluster restrictions were tested at 2, 3, and 4 maximum clusters and then compared to find the optimal number of clusters within the data. The median between cluster means within the parental generation at 21 DAT at each herbicide rate were used to define the resistance threshold for binary classes at that specific herbicide rate. Plants in F1 and F2 generations that received an injury score below the median of the parental cluster injury means for that herbicide rate were classified as ‘resistant’ and assigned ‘1’. Scores higher than the median were ‘susceptible’ and assigned ‘0’. After binary classes were assigned to each plant, logistic regression was used to indicate the probability of resistance for each unique combination of plant generation, herbicide rate, and DAT (PROC LOGISTIC). Any comparisons of the probability of resistance were made using LSMEANS and LSMESTIMATE.

2.6.2 Chi-Square Goodness-of-fit Test of F1 and F2 Generations

The chi-square GOF test was used to test the hypothesis that a small number of major segregating genes are primarily responsible for the non-target-site resistance (NTSR). The three models used for the chi-square are derived from ratios based on one, two, and three major loci containing the resistance trait. Important assumptions included in the models were; 1) the susceptible parent (CFS) is homozygous recessive at the NTSR
locus or loci, 2) the resistant parent (HCR\textsubscript{3}) is heterozygous at the NTSR locus or loci, and 3) at least one dominant allele at each loci must be present for the plant to confer herbicide resistance. Data from the binary classes was used to designate the number of observed surviving plants in the chi-square test. The null hypotheses (H\textsubscript{0}) of the F\textsubscript{1} generation failed to be rejected when the number of surviving plants was not significantly different than the expected number of surviving plants of the F\textsubscript{1} generation based on the R:S segregation for one (1:1), two (1:3), or three (1:7) loci. The null hypotheses (H\textsubscript{0}) of the F\textsubscript{2} generation failed to be rejected when the number of surviving plants was not significantly different than the expected number of surviving plants of the F\textsubscript{2} generation based on the R:S segregation of one (5:12), two (29:160), or three (185:2,306) loci. The null hypotheses were tested at a significance level of \( \alpha = 0.05 \).

3. Results

3.1 Parent Generation Mesotrione Dose Response

The parent waterhemp populations (CFS and HCR\textsubscript{3}) were treated with a range of mesotrione rates to determine the dose responses for susceptible and HPPD resistant waterhemp biotypes. The effect of populations, mesotrione rate, and DAT were confirmed to be statistically different while the interaction of population, rate, and DAT was not statistically different indicating that the differences between the populations were significant and consistent across rates. (Table 1). The mean percent damage for the CFS parent was always higher than the HCR\textsubscript{3} parent (Table 2). The susceptible response of the CFS parent population and the increase in susceptibility of the HCR\textsubscript{3} parent population were both consistent across rates (Figure 2). When comparing populations at each herbicide rate 21 DAT, the mean percent damage for the HCR\textsubscript{3} parent population
remained significantly different than the CFS parent population up to the 8x rate (Table 3), validating that the HCR$_3$ parent population is highly resistant. The mean percent damage of the CFS parent population across all herbicide rates 21 DAT was not statistically different (p value = 0.8131), confirming the CFS population to be highly susceptible, even at the lowest herbicide rate. The high level of susceptibility of the CFS parent population to the mesotrione prevented an accurate estimation of a median lethal dose (LD$_{50}$) value for the CFS population, establishing accurate dose-response curves, and calculating the susceptibility fold change in the HCR$_3$ population.

Previously, the same resistant waterhemp population (before recurrent selection was implemented) was characterized to be 8-fold less sensitive to mesotrione than a susceptible waterhemp population [18]. McMullan and Green (2011) performed the experiment in the field, while our study was performed in the greenhouse. The HPPD-resistant waterhemp population from Illinois was calculated to be 10- or 35- fold less susceptible, depending on the HPPD-susceptible waterhemp population used for comparison [19]. Interestingly, the 8-fold resistance level described by McMullan and Green (2011) and the comparable 10-fold resistance from Hausman et al. (2011) were both characterized to be resistant to atrazine as well as mesotrione. Despite slight differences in methodology, waterhemp response to mesotrione were similar between this and the two previous studies. These comparisons help validate the expected response and the use of HCR$_3$ as the resistant parent population by showing a similar responses of mesotrione resistant waterhemp among different resistant populations.
3.2 Characterizing the Inheritance of HPPD-Resistance

3.2.1 Frequency of resistance across plant generations

Initially no significant differences were found comparing the probability of resistance of the F₁ and the F₂ generations across all four mesotrione rates 21 DAT when comparing cumulative data of reciprocal crosses within each generation (Figure 3A and Table 4). When the data was separated by reciprocal crosses at each rate (Figure 3B and Table 5), significant differences between the F₁ and F₂ generations were found in HCR₃-♀ at the 2x and 4x rate (p-value = 0.0017 and 0.0084, respectively).

Finding no significant differences between generations at any herbicide rate may suggest the HPPD-resistance trait is relatively stable regardless of consistent selection pressure from the herbicide. Without eliminating the susceptible phenotypes from the population gene pool, there is no driving force behind the accumulation of resistance alleles needed for plants to express a resistant phenotype sufficient for survival [25].

3.2.2 Nuclear Inheritance

In order to determine an effect of maternal inheritance we compared the differences in the probability of resistance between the reciprocal crosses within in the F₁ and F₂ generations using the odds ratio (OR) [26]. In this instance, the OR is the ratio of the probability of resistance of HCR₃-♀ to the probability of resistance of HCR₃-♂. If the OR is greater than one, the probability of resistance of HCR₃-♀ is higher, if the OR is less than one, the probability of resistance of HCR₃-♂ is higher, or if the OR is equal to one the probabilities of resistance of HCR₃-♀ and HCR₃-♂ are the same.

The responses of reciprocal crosses for the 1x, 2x, and 4x herbicide rates in the F₁ generation were all found to be significantly different (p-value = 0.0246, 0.0161, and
0.0052, respectively) (Table 6 and Figure 3B). In all three statistically significant herbicide rate responses, the probability of resistance of HCR3-♀ is consistently higher than HCR3-♂, given that the ORs for all three rates are greater than one. The responses of reciprocal crosses in the 0.5x rate for the F1 generation and all four rates in the F2 generation were not significantly different.

When a trait is maternally inherited, there is a greater contribution of the female plant phenotype than expected in the ensuing progeny [27]. Cytoplasmic inheritance is a mechanism of maternal inheritance that can contribute to heritable variation of offspring in quantitative traits. In this experiment, that would mean if the HPPD resistance trait is from cytoplasmic inheritance, we should expect to see differences in proportions of the resistant phenotype in progeny from the resistant female parents (HCR3-♀) and the susceptible female parent (HCR3-♂). Even though we did find a significant difference between the reciprocal crosses in the F1 generation, we still found a comparable level of HPPD resistance in the progeny of the susceptible female parent (HCR3-♂) supporting the hypothesis that the resistance trait is due to nuclear inheritance as opposed to cytoplasmic inheritance. No significant differences comparing the reciprocal crosses for the F2 generation were observed, which further supports the hypothesis of nuclear inheritance for HPPD resistance. The results are consistent with the previous report [21].

3.2.3 Characteristics of Resistance Loci and Alleles

Traits that are easily classified into separate classes (qualitative traits) are typically controlled by fewer genes [28]. More complex traits that can be measured across a spectrum of responses (quantitative traits) tend to have more genes involved in trait expression. Our initial analysis looked for a distinct segregation of classes within the
data. Early cluster analysis revealed no consistency for the number of natural clusters derived within each generation at each herbicide rate. After restricting the maximum number of potential clusters used by the cluster analysis, the optimal number of for each generation by rate combination was 2 clusters. Clear differences of herbicide response between the parent populations at each rate were observed (Figure 4). More important was the observed continuous distribution of herbicide response in the F₁ and F₂ generation (particularly at the 0.5x rate); this suggests the resistance is a quantitative trait (Figure 4A). As mesotrione rate increases, the symmetry of the F₁ and F₂ herbicide responses decreases (Figures 4B, 4C, and 4D). The decreasing symmetry seen in the herbicide response of the F₁ and F₂ generations can be associated with the change in the herbicide response of the resistant parent population. The herbicide response of the resistant parent population becomes less distinct as mesotrione rate increases suggesting an increase in relative sensitivity. The change in the herbicide response of the resistant parent suggests that there should be a similar change expected in the herbicide response of the F₁ and F₂ generations. This trend in herbicide response as mesotrione rate increases is supported by the consistent decrease in the probability of mesotrione resistance of the parent resistant population, F₁ generation, and F₂ generation (Figure 3A). The dynamic response found in the in the F₁ and F₂ generations are a strong indication that HPPD resistance is a quantitative trait, supporting the hypothesis that HPPD-resistance trait in waterhemp is polygenic [28].

The herbicide response of the F₁ and F₂ generations were continuously distributed preventing the opportunity to clearly assign a herbicide injury response as the resistance threshold, which would help distinguish plant response as resistant or susceptible.
this reason we used the herbicide response of the parent populations at each rate to determine the herbicide injury response we would use as the resistance threshold for each respective herbicide rate. Based on the cluster analysis of the parent populations, the resistance thresholds we used for the 0.5x, 1x, 2x, and 4x rate were 55.17%, 55.55%, 61.22%, and 65.16%, respectively. After sorting the data as resistant or susceptible using the parent population resistance threshold we found a similar R:S frequency as the cluster analysis.

Creating resistance thresholds for each rate and sorting the data as resistant or susceptible allowed us to perform the chi-square GOF testing for the presence of major segregating genes for the HPPD resistance trait. The segregation ratios for the F1 generation response to mesotrione rates at 21 DAT were successfully fitted to all three models: The 0.5x rate fit the single gene model, the 1x rate fit the two gene model and the 2x and 4x rate both fit the three gene model, initially suggesting the resistance trait to be polygenic (Table 7). Due to the assumption of a heterozygous resistant parent, there are more potential genotypes in the F1 seed than if the resistant parent was homozygous dominant. In order to test segregation ratios in the F2 generation we first had to determine all of the potential genotypic outcomes that could occur as a result of inter-mating the F1 generation and then determine the number of potential genotypes that would express resistance in the F2 seed according to our assumption that at least one dominant allele must be present at each loci for resistance to occur. After determining all potential genotypes, we found the segregation ratios at each mesotrione rate for the F2 generation at 21 DAT successfully fit all three models: The 0.5x rate fit the single gene model, the 1x rate fit the two gene model and the 2x and 4x rate both fit the three gene model,
further suggesting the resistance trait to be polygenic (Table 8). Not only were all models successfully fitted in both generations but the model successfully fit at each rate in the F2 generation corresponded to the same successful fit for the F1 generation, adding support our original assumptions. The major observation we found in the chi-square analyses of the F1 and F2 generations was as mesotrione rate increased, the number of genes required to fit the model increased as well. According to our assumptions, this means in order for waterhemp to be resistant at higher mesotrione rates, more major genes containing at least one dominant resistance allele must be expressed. Furthermore, the successful fit of models using two and three genes supports the hypothesis that HPPD resistance in waterhemp is polygenic.

4. Discussion

When characterizing the inheritance of a trait, it is important to establish that the parent populations are homogeneous for the trait. Collecting seed from a putative HPPD-resistant field-grown waterhemp population and recognizing the dioecious sexual habit, the homogeneity of the resistance trait within the population must be questioned. Establishing multiple rounds of phenotypic recurrent selection based on herbicide response increases the homogeneity of the resistance trait in subsequent generations. However, even though there would be an increase in the consistency of the resistance phenotype in the population, there is no assurance of a favorable shift towards homozygosity of the resistance trait. The obligate outcrossing, dioecious nature of waterhemp forces the consistent reshuffling of genes for each generation creating an enormous number of possible gene combinations within a gene pool [28].
Huffman et al. (2015) used only one round of phenotypic selection based on herbicide response in their resistant parent population while we used multiple rounds of phenotypic recurrent selection. They calculated the degree of dominance of mesotrione resistance based on the formula provided by Stone (1968) using the LD$_{50}$ of the two parent populations and the F$_1$ generation. We were unable to perform the same calculation because the high level susceptibility of our susceptible parent population (CFS) prevented an accurate calculation of the susceptible parent LD$_{50}$. Based on the degree of dominance calculation, Huffman et al. (2015) estimated mesotrione resistance in waterhemp to be incompletely recessive; F$_1$ plants were slightly more resistant than the susceptible parent [29]. They performed their segregation analysis at mesotrione rates where resistance was functionally dominant. Huffman et al. (2015) eventually concluded HPPD-resistance in waterhemp is likely polygenic and based on our chi-square analyses, we refine and support their conclusion.

When dealing with multiple genes, the interaction of alleles between genes as an alternative explanation of the expected phenotypic response must be considered. Epistasis occurs when the phenotypic expression of one gene changes due to the presence of another gene, thus changing the expected trait segregation ratios [27]. Huffman et al. (2015) proposed the possibilities of epigenetic effects or conditional dominance in response to an environmental condition, as explanations for the resistance segregation ratios.

The potential conditional effect of mesotrione rate on the number of major genes involved with the expression of HPPD-resistance has applicable repercussions to weed management in production maize fields. The use of HPPD-inhibiting herbicide in the
field at lower-than-recommended rates could favor the evolution of stronger HPPD-resistance. In theory, if waterhemp plants only need one major resistance gene to survive lower herbicide rates, that provides the opportunity for plants containing different major resistance genes to recombine and “stack” a number of major resistance genes thus increasing the resistance in the population to higher herbicide doses [30–32]. *Lolium rigidum* (Gaud.) evolved polygenic resistance after repeated exposure to low herbicide doses [33]. Similar to waterhemp, *L. rigidum* is an obligate outcrosser and exhibits enhanced herbicide metabolism to confer herbicide resistance.

In conclusion, we have validated the inheritance of mesotrione resistance in an Iowa waterhemp population to be complex. Our findings suggest the mesotrione resistance trait to have nuclear polygenic inheritance; the dioecious nature of waterhemp prevented the simple characterization of dominance. Our results suggest at least one dominant allele at each participating major loci is needed to confer resistance and more major genes contributing to the resistance allows plant survival at higher mesotrione rates. Heritability studies of herbicide resistance can help form a better understanding of the spread and development of herbicide resistance and stress the need for better management strategies to combat the rapid expansion of herbicide resistance.

**References**


Figure 1. (A) Parent cross (2013) tent layout (6.1m x 2.1m). Diagonal shading identifies matching quadrants with plants of the same gender. (B) Parent cross (2013) plant distribution within one quadrant of the tent. 50 plants, represented as dots, were evenly distributed in a 5 x 10 grid.
Figure 2. The mean percent damage response of susceptible and resistant waterhemp parent populations (CFS and HCR$_3$, respectively) across five mesotrione rates evaluated 21 days after treatment in the greenhouse.
Figure 3. The probability of mesotrione resistance in three waterhemp generations 21 days after treatment. HCR₃ and CFS represent the resistant and susceptible parent populations, respectively. (A) Cumulative data of the F₁ and F₂ generations. (B) Data from the F₁ and F₂ generations separated by reciprocal crosses. HCR₃-♀ designates the F₁ families derived from parent cross (HCR₃-♀ x CFS-♂) and HCR₃-♂ designates the F₁ families derived from the parent cross (CFS-♀ x HCR₃-♂).
Figure 4. The response of three waterhemp generations to four mesotrione rates, where the 1x is 105 g ai ha⁻¹, 21 days after treatment. HCR₃ and CFS represent the resistant and susceptible parent populations, respectively. HCR₃-♀ designates the families in both F₁ and F₂ derived from parent cross (HCR₃-♀ x CFS-♂). HCR₃-♂ designates the families in both F₁ and F₂ derived from the parent cross (CFS-♀ x HCR₃-♂).
Figure 4. continued
Figure 4. continued
Figure 4. continued
Tables

Table 1. Tests of fixed effects in the susceptible and resistant waterhemp parent populations (CFS and HCR₃, respectively). The populations were treated with five mesotrione rates (53, 105, 211, 421, and 843 g ai ha⁻¹) and evaluated at 7, 14, and 21 days after treatment.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DFᵃ</th>
<th>Den DFᵇ</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATᶜ</td>
<td>2</td>
<td>421</td>
<td>118.06</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Population</td>
<td>1</td>
<td>90</td>
<td>226.42</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Population*DAT</td>
<td>2</td>
<td>421</td>
<td>39.92</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Rate</td>
<td>4</td>
<td>90</td>
<td>4.86</td>
<td>0.0014</td>
</tr>
<tr>
<td>Rate*DAT</td>
<td>8</td>
<td>421</td>
<td>2.6</td>
<td>0.0089</td>
</tr>
<tr>
<td>Population*Rate</td>
<td>4</td>
<td>90</td>
<td>0.56</td>
<td>0.6921</td>
</tr>
<tr>
<td>Pop<em>Rate</em>DAT</td>
<td>8</td>
<td>421</td>
<td>0.49</td>
<td>0.8603</td>
</tr>
</tbody>
</table>

ᵃ = degrees of freedom in the numerator
ᵇ = degrees of freedom in the denominator
ᶜ = days after treatment

Table 2. Summary of the mean percent tissue damage of the susceptible and resistant waterhemp parent populations (CFS and HCR₃, respectively) at five mesotrione rates, where the 1x is 105 g ai ha⁻¹ evaluated 21 days after treatment

<table>
<thead>
<tr>
<th>Rate</th>
<th>Family</th>
<th>DFᵃ</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Mean Damageᵇ</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5x</td>
<td>CFS³</td>
<td>421</td>
<td>5.79</td>
<td>&lt; 0.0001</td>
<td>0.97</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>HCR</td>
<td>421</td>
<td>-4.66</td>
<td>&lt; 0.0001</td>
<td>0.22</td>
<td>0.046</td>
</tr>
<tr>
<td>1x</td>
<td>CFS³</td>
<td>421</td>
<td>5.07</td>
<td>&lt; 0.0001</td>
<td>0.98</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>HCR</td>
<td>421</td>
<td>-2.98</td>
<td>0.0031</td>
<td>0.33</td>
<td>0.054</td>
</tr>
<tr>
<td>2x</td>
<td>CFS³</td>
<td>421</td>
<td>4.26</td>
<td>&lt; 0.0001</td>
<td>0.99</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>HCR</td>
<td>421</td>
<td>-0.64</td>
<td>0.5216</td>
<td>0.46</td>
<td>0.061</td>
</tr>
<tr>
<td>4x</td>
<td>CFS³</td>
<td>421</td>
<td>4.48</td>
<td>&lt; 0.0001</td>
<td>0.99</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>HCR</td>
<td>421</td>
<td>2.63</td>
<td>0.0088</td>
<td>0.67</td>
<td>0.059</td>
</tr>
<tr>
<td>8x</td>
<td>CFS³</td>
<td>421</td>
<td>4.46</td>
<td>&lt; 0.0001</td>
<td>0.99</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>HCR</td>
<td>421</td>
<td>5.08</td>
<td>&lt; 0.0001</td>
<td>0.83</td>
<td>0.044</td>
</tr>
</tbody>
</table>

ᵃ = degrees of freedom
ᵇ = average percent tissue damage
Table 3. Comparing the mean percent tissue damage of the susceptible and resistant waterhemp parent populations (CFS and HCR₃, respectively) at five mesotrione rates, where the 1x is 105 g ai ha⁻¹, 21 days after treatment.

| Rate | DF ¹ | t Value | Pr > |t| |
|------|------|---------|------|---|
| 0.5x | 421  | 7.17    | < 0.0001 |
| 1x   | 421  | 5.74    | < 0.0001 |
| 2x   | 421  | 4.29    | < 0.0001 |
| 4x   | 421  | 3.68    | 0.0003  |
| 8x   | 421  | 2.81    | 0.0052  |

¹ = degrees of freedom

Table 4. Comparing the probability of resistance of the F₁ and F₂ waterhemp generations at four mesotrione rates, where the 1x is 105 g ai ha⁻¹ evaluated 21 days after treatment.

| Rate | Estimate | Standard Error | z Value | Pr > |z| | Odds Ratio ¹ |
|------|----------|----------------|---------|------|---|--------------|
| 0.5x | 0.218    | 0.1849         | 1.18    | 0.2383 | 1.2436 |
| 1x   | -0.09908 | 0.218          | -0.45   | 0.6495 | 0.9057 |
| 2x   | 0.2922   | 0.279          | 1.05    | 0.295  | 1.3393 |
| 4x   | 0.2073   | 0.3451         | 0.6     | 0.548  | 1.2303 |

¹ The odds ratio = odds of resistance in F₁ / odds of resistance in F₂. If ratio > 1, the probability of resistance is higher in the F₁, if ratio < 1 the probability of resistance is higher in the F₂, if ratio = 1 the probability of resistance is the same.
Table 5. Comparing the probability of resistance of the F₁ and F₂ waterhemp generations within HCR₃-♀ and HCR₃-♂ at four mesotrione rates, where the 1x is 105 g ai ha⁻¹, 21 days after treatment. HCR₃-♀ designates the families derived from parent cross (HCR₃-♀ x CFS-♂). HCR₃-♂ designates the families derived from the parent cross (CFS-♀ x HCR₃-♂).

<table>
<thead>
<tr>
<th>Tent</th>
<th>Rate</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>z Value</th>
<th>Pr &gt;</th>
<th>Odds Ratio *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5x</td>
<td>0.3295</td>
<td>0.2419</td>
<td>1.36</td>
<td>0.1732</td>
<td>1.3902</td>
</tr>
<tr>
<td>HCR₃-♀</td>
<td>1x</td>
<td>0.2945</td>
<td>0.2593</td>
<td>1.14</td>
<td>0.256</td>
<td>1.3425</td>
</tr>
<tr>
<td></td>
<td>2x</td>
<td>0.9873</td>
<td>0.3143</td>
<td>3.14</td>
<td>0.0017</td>
<td>2.684</td>
</tr>
<tr>
<td></td>
<td>4x</td>
<td>0.8393</td>
<td>0.3185</td>
<td>2.64</td>
<td>0.0084</td>
<td>2.3148</td>
</tr>
<tr>
<td>HCR₃-♂</td>
<td>0.5x</td>
<td>0.1065</td>
<td>0.2796</td>
<td>0.38</td>
<td>0.7032</td>
<td>1.1124</td>
</tr>
<tr>
<td></td>
<td>1x</td>
<td>-0.4927</td>
<td>0.3506</td>
<td>-1.41</td>
<td>0.16</td>
<td>0.611</td>
</tr>
<tr>
<td></td>
<td>2x</td>
<td>-0.403</td>
<td>0.461</td>
<td>-0.87</td>
<td>0.382</td>
<td>0.6683</td>
</tr>
<tr>
<td></td>
<td>4x</td>
<td>-0.4247</td>
<td>0.6122</td>
<td>-0.69</td>
<td>0.4878</td>
<td>0.6539</td>
</tr>
</tbody>
</table>

* The odds ratio = odds of resistance in F₁ / odds of resistance in F₂. If ratio > 1, the probability of resistance is higher in the F₁, if ratio < 1 the probability of resistance is higher in the F₂, if ratio = 1 the probability of resistance is the same.

Table 6. Comparing the probability of resistance of the HCR₃-♀ and HCR₃-♂ within the F₁ and F₂ waterhemp generations at four mesotrione rates, where the 1x is 105 g ai ha⁻¹, 21 days after treatment. HCR₃-♀ designates the families derived from parent cross (HCR₃-♀ x CFS-♂). HCR₃-♂ designates the families derived from the parent cross (CFS-♀ x HCR₃-♂).

<table>
<thead>
<tr>
<th>Generation</th>
<th>Rate</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>z Value</th>
<th>Pr &gt;</th>
<th>Odds Ratio *</th>
</tr>
</thead>
<tbody>
<tr>
<td>F₁</td>
<td>0.5x</td>
<td>0.537</td>
<td>0.3206</td>
<td>1.67</td>
<td>0.094</td>
<td>1.7108</td>
</tr>
<tr>
<td></td>
<td>1x</td>
<td>0.8592</td>
<td>0.3823</td>
<td>2.25</td>
<td>0.0246</td>
<td>2.3613</td>
</tr>
<tr>
<td></td>
<td>2x</td>
<td>1.1623</td>
<td>0.4831</td>
<td>2.41</td>
<td>0.0161</td>
<td>3.1973</td>
</tr>
<tr>
<td></td>
<td>4x</td>
<td>1.6963</td>
<td>0.6067</td>
<td>2.8</td>
<td>0.0052</td>
<td>5.4528</td>
</tr>
<tr>
<td>F₂</td>
<td>0.5x</td>
<td>0.314</td>
<td>0.1841</td>
<td>1.71</td>
<td>0.0881</td>
<td>1.3689</td>
</tr>
<tr>
<td></td>
<td>1x</td>
<td>0.07197</td>
<td>0.2098</td>
<td>0.34</td>
<td>0.7315</td>
<td>1.0746</td>
</tr>
<tr>
<td></td>
<td>2x</td>
<td>-0.228</td>
<td>0.2791</td>
<td>-0.82</td>
<td>0.414</td>
<td>0.7931</td>
</tr>
<tr>
<td></td>
<td>4x</td>
<td>0.4322</td>
<td>0.3289</td>
<td>1.31</td>
<td>0.1887</td>
<td>1.5407</td>
</tr>
</tbody>
</table>

* The odds ratio = odds of resistance in HCR₃-♀ / odds of resistance in HCR₃-♂. If ratio > 1, the probability of resistance is higher in the HCR₃-♀, if ratio < 1 the probability of resistance is higher in the HCR₃-♂, if ratio = 1 the probability of resistance is the same.
Table 7. Chi-square analysis for goodness-of-fit (GOF) test of the observed segregation of mesotrione resistance in the F₁ waterhemp generation at four rates of mesotrione, where 1x is 105 g ai ha⁻¹ evaluated 21 days after treatment.

<table>
<thead>
<tr>
<th>Rate</th>
<th>Number of F₁ Plants</th>
<th>GOF Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Total</td>
</tr>
<tr>
<td>0.5x</td>
<td>83</td>
<td>162</td>
</tr>
<tr>
<td>1x</td>
<td>43</td>
<td>165</td>
</tr>
<tr>
<td>2x</td>
<td>24</td>
<td>153</td>
</tr>
<tr>
<td>4x</td>
<td>23</td>
<td>162</td>
</tr>
</tbody>
</table>

ᵃ Expected values are calculated from a 1:1 resistant:susceptible (R:S) ratio.
ᵇ Expected values are calculated from a 1:3 (R:S) ratio.
ᶜ Expected values are calculated from a 1:7 (R:S) ratio.

Table 8. Chi-square analysis for goodness of fit (GOF) test of the observed segregation of mesotrione resistance in the F₂ waterhemp generation at four rates of mesotrione, where 1x is 105 g ai ha⁻¹ evaluated 21 days after treatment.

<table>
<thead>
<tr>
<th>Rate</th>
<th>Number of F₂ Plants</th>
<th>GOF Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Total</td>
</tr>
<tr>
<td>0.5x</td>
<td>221</td>
<td>488</td>
</tr>
<tr>
<td>1x</td>
<td>83</td>
<td>476</td>
</tr>
<tr>
<td>2x</td>
<td>41</td>
<td>471</td>
</tr>
<tr>
<td>4x</td>
<td>37</td>
<td>480</td>
</tr>
</tbody>
</table>

ᵃ Expected values are calculated from a 5:12 resistant:susceptible (R:S) ratio
ᵇ Expected values are calculated from a 29:160 (R:S) ratio
ᶜ Expected values are calculated from a 185:2306 (R:S) ratio
CHAPTER 3: AN RNA-SEQ DE NOVO TRANSCRIPTOME ASSEMBLY OF AMARANTHUS TUBERCU LATUS AND ANALYSES OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS RELATED TO 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE INHIBITOR HERBICIDE RESISTANCE

A paper to be submitted to BMC Genomics

Daniel Kohlhase, Mike Owen, Michelle Graham

Abstract

Waterhemp (Amaranthus tuberculatus) is a problem weed commonly found in the Midwestern United States that can cause yield losses up to 74% and 56% in maize (Zea mays L.) and soybean [Glycine max (L.) Merr.], respectively. One advantageous adaptive characteristic found in waterhemp is the ability to evolve herbicide resistance. Waterhemp populations are resistant to six herbicide sites of action. In 2011, p-hydroxyphenylpyruvate-dioxygenase (HPPD, EC 1.13.11.27) inhibitor herbicide resistance was reported in two waterhemp populations. Since the discovery of HPPD-herbicide resistance, studies have identified the mechanism of resistance and described the inheritance of the herbicide resistance trait. Currently, there are no studies that have examined the transcriptomic expression response of HPPD-herbicide resistance in waterhemp. However, the genomic resources of waterhemp are limited; therefore we conducted an RNA-sequencing (RNA-seq) de novo transcriptome assembly of waterhemp. We treated and mock-treated two waterhemp populations (HPPD-herbicide resistant and susceptible) and collected leaf samples at three, six, twelve, and twenty-four
hours after treatment (HAT). We performed a de novo transcriptome assembly using all sample sequences to better represent the waterhemp transcriptome as well as allowing us to identify transcripts specific to a genotype, treatment, or time point. Our results indicate that the response of HPPD-herbicide resistant and susceptible waterhemp genotypes to HPPD-inhibiting herbicide is very rapid and established as soon as three HAT. Furthermore, the differentially expressed transcripts expressed within a genotype in response to herbicide treatment showed little overlap between genotypes. Unique Arabidopsis thaliana identifiers, derived from A. thaliana homologs of unique differentially expressed transcripts in waterhemp, also suggest the possibility of overlapping gene networks in response to other herbicides. We have made available the raw sequences, and assembled sequences with complete annotations for continued use by the weed science community.

**Introduction**

Over the past 30 years waterhemp (Amaranthus tuberculatus (Moq.) J.D. Sauer) has evolved into a major problem weed species in agriculture across the Midwest United States [1]. If not properly managed, fields infested with waterhemp can suffer yield losses up to 74% in maize (Zea mays L.) and 56% in soybean [Glycine max (L.) Merr.] [2,3]. Waterhemp is native to the Midwest United States and is dioecious; the male and female reproductive structures are on separate plants. The dioecious nature of waterhemp forces plants to outcross, resulting in endless opportunities for genetic recombination. Obligatory outcrossing facilitates the movement of ecologically valuable traits, such as herbicide resistance, among and between waterhemp populations. Additional biological
traits that contribute to the weediness of waterhemp include prolific seed production [4], extended and opportunistic germination [5], and rapid growth rate [6].

Herbicides are the most important tool in weed management for most crop production systems in many parts of the world [7]. One of the first documented cases of evolved herbicide resistance in weeds was reported in 1970 and since then, the number of unique cases (a case is defined as an individual species by specific herbicide site of action) has grown to a currently reported 471 globally [8,9]. Mesotrione (2-(4-Mesyl-2-nitrobenzoyl)-1,3-cyclohexanedione, Herbicide Group (HG) 27) is a selective herbicide that inhibits 4-hydroxyphenylpyruvate dioxygenase (HPPD, EC 1.13.11.27) and is commonly used for broadleaf control in maize [10]. HPPD converts 4-hydroxyphenylpyruvate (4-HPP) to homogentisate (2,5-dihydroxyphenylacetate; HGA), which is an important precursor in carotenoid biosynthesis. The herbicidal activity of mesotrione is characterized by the bleaching of new tissue followed by tissue necrosis. In 2011, two waterhemp populations with evolved resistance to HPPD-inhibiting herbicides were discovered in seed maize fields in Iowa and Illinois [11,12]. The mechanism of resistance for the Illinois waterhemp population was determined to be by the metabolism of the parent mesotrione molecule to non-herbicidal metabolites and was reportedly attributable to increased cytochrome P450 monooxygenase (Cyt P450, EC 1.14.14.1) activity [13]. Additionally, the inheritance of the HPPD resistance in this waterhemp population was reported to be “complex” [14]. Despite a genetic examination of the Illinois HPPD-resistant waterhemp population, gene expression in HPPD-resistant waterhemp has not been described. Characterizing gene expression in putative HPPD-resistant waterhemp populations could help identify major genes contributing to
resistance and may provide insight into managing the evolution of resistance to other herbicides in waterhemp and possibly other weed species.

In recent years, advances in genomic sequencing technology have created more opportunities to study the genomics of non-model organisms [15]. Due to a lack of weed-related genomic resources, Lee et al. [15] sampled 43 million base pairs of the waterhemp genome using 454 pyrosequencing (Roche Sequencing, Pleasanton, CA, USA). While this sequencing approach covered less than 10% of the waterhemp genome, it demonstrated that cutting-edge sequencing technology could be applied to weed species. Riggins et al. [16] used 454 pyrosequencing to analyze the waterhemp transcriptome. To maximize transcriptome coverage, the study pooled RNA samples from different individuals, sexes, tissues, life stages, herbicide treatments and cold stress. These studies contributed to a better understanding of the waterhemp genome and provided sequence-based details for key enzymes targeted by herbicides and potentially prone to herbicide resistance evolution [16]. However, these studies failed to address the genes and gene networks that regulate susceptibility and tolerance to herbicides. Since these initial studies, RNA-sequencing (RNA-seq) has become the standard method for transcriptome analyses for species lacking genomic resources.

The increasing prominence of waterhemp as an economically important and ubiquitous weed in Midwestern United States and the demonstrated ability to evolve resistance to herbicides makes this species an important model for studying herbicide resistance evolution in weeds. Unfortunately, genomic knowledge of waterhemp is limited. Here we report on the sequencing of the waterhemp transcriptome using high throughput RNA-seq technology. This study identifies the genes and gene networks
responding to the HPPD-inhibiting herbicide mesotrione in susceptible and resistant waterhemp genotypes over a 24 hour exposure time course. In addition, our study provides a publicly available sequenced-based platform for the weed science community to study this agronomically important weed.

**Materials and Methods**

**Tissue Collection**

Two waterhemp populations with different susceptibility phenotypes to HPPD-inhibiting herbicides (susceptible and resistant) were selected. The susceptible waterhemp population was collected from the Curtis Farm at Iowa State University (Ames, IA, USA) in 2006. The resistant population was from Tarnov, Nebraska (USA) and was collected in 2014 from a field with a history of seed maize production after receiving reports of waterhemp surviving multiple applications of mesotrione.

Each genotype was planted in 40 individual 15.2 cm round pots using a 4:1 mixture of Sunshine Mix #1/LC1 (Sun Gro Horticulture, Agawam, MA, USA) to sand ratio, respectively. We added 1 tsp of Osmocote Flower Food Granules (14-14-14) (The Scotts Miracle-Gro Company, Marysville, OH, USA) to each pot at the time of planting. Plants were grown in a greenhouse set to 24°C with a 14 hour photoperiod supplemented by high pressure sodium bulbs. Plants were watered every other day. After two weeks, seedlings were thinned to 3 plants per pot. Each plant within each pot was randomly assigned a label of A, B, or C. The pots were placed in the greenhouse in a randomized block design.

When plants reached a minimum height of 7.6 cm, they were treated with mesotrione applied in a CO₂ powered spray chamber equipped with TeeJet® 80015EVS
nozzles (Spraying Systems Co., Wheaton, IL, USA) at a carrier volume of 191.76 L ha\textsuperscript{-1}. Half of each population (20 pots of each genotype) was treated with 105.36 g ai ha\textsuperscript{-1} of mesotrione, 1% (v/v) crop oil concentrate (COC), 2.5% (v/v), urea ammonium nitrate solution (UAN) and water. The other half was treated with water, representing a mock treatment. The plants were then returned to the greenhouse into 4 blocks; the blocks were separated by treatment and genotype. In a previous experiment we evaluated the phenotypic response of plants that were untreated, mock treated, and treated with 1% (v/v) crop oil concentrate (COC), 2.5% (v/v), urea ammonium nitrate solution (UAN) and water (data not provided). We found no difference in phenotypic response. Therefore, to help reduce the cost of sequencing, we chose to use just the mock treatment to help reduce the number of sequenced samples.

Each genotype within a treatment was separated into 4 groups of 5 pots. The 4 groups were randomly assigned a time point of 3, 6, 12, or 24 HAT. Within each time point 4 pots were labeled 1-4. The fifth pot was used as a control for verification of the phenotypic response and was also used as a buffer against greenhouse variation in the bench space adjacent to the wall. Leaf tissue from each plant within the four labeled pots was collected 3, 6, 12, and 24 HAT. A given plant was only sampled at one time point. The four youngest fully-developed leaves of each plant were excised at the base of each leaf, placed in a 50 mL Falcon\textregistered tube (Thermo Fisher Scientific, Waltham, MA, USA), flash frozen in liquid nitrogen, and then maintained at -80\degree C. Tissues from individual plant samples were stored in a separate Falcon\textregistered tube. Plants continued to grow for 3 weeks after treatment to verify the phenotypic response to mesotrione.
**RNA Isolation**

Frozen tissue in the 50 mL Falcon® tubes was crushed by inverting an 11.11 cm pestle, dipped in liquid nitrogen, into the tubes. Crushing the leaf samples within a Falcon® tube adequately mixed the tissue from an individual plant providing a more homogeneous collection of leaves from each plant. One full microspatula scoop (approximately 100 mg) of crushed frozen tissue from each Falcon® tube was added to a 2 mL Safe-Lock™ microcentrifuge tube (Eppendorf, Hamburg, Germany) kept on dry ice with a 3 mm tungsten carbide bead. Prepared microtubes were placed in TissueLyser Adapter sets precooled at -80°C and then processed in a Qiagen TissueLyser II (Qiagen, Valencia, CA, USA) for 1 minute at 30 Hz. RNA extraction was performed as recommended by the manufacturer using the RNeasy® Plant Mini Kit (Qiagen, Valencia, CA, USA). To check for RNA concentration and quality, a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was used.

Before the samples were DNAsed to remove genomic DNA contamination, RNA samples from plants growing in the same pot were pooled together. 6 µg of pooled RNA (2 µg of RNA per sample) for a 50 µL total reaction volume. The Ambion® TURBO DNA-free™ Kit (Thermo Fisher Scientific, Waltham, MA, USA) was used to remove DNA contamination and then immediately afterwards the samples were further purified using the RNeasy® MinElute® Cleanup Kit following the manufacturer’s recommendation. RNA concentration and quality of the samples was checked using the NanoDrop™ 1000 Spectrophotometer. In addition, the 24 HAT samples were processed on a QIAxcel (Qiagen, Valencia, CA, USA) for another sample quality check. The 24 HAT time point was the only time point used because of the 4 time points, it had the
longest exposure to the herbicide and thus was most indicative of RNA quality and quantity when checking for signs of RNA degradation.

**RNA-Seq and de novo Transcriptome Assembly**

The extracted RNA was sequenced by the Iowa State University DNA Facility using the Illumina HiSeq 2500 (Illumina, Inc., San Diego, CA, USA) platform. Sequences were generated in High Output Mode with 100 base pair read length and paired-end sequencing. The paired-end protocol, sequencing the RNA from both directions of the strand, enables better transcriptome coverage. 48 samples were run on one eight-lane flow cell, six samples per lane. Each lane contained three samples of each treatment (herbicide or mock), of one waterhemp genotype (resistant or susceptible), at one time point (3, 6, 12, or 24 HAT).

The programs Scythe (UC Davis Bioinformatics Core, https://github.com/ucdavis-bioinformatics) and Sickle (UC Davis Bioinformatics Core, https://github.com/ucdavis-bioinformatics) were used to remove sequencing artifacts, low quality bases (q<20) and short reads (l<50) from all 48 sequenced samples. Trinity (version 2.0.6, [17]) was used to produce multiple de novo transcriptome assemblies. Three separate assemblies (versions 1-3) were made using kmer lengths of 25, 29, and 32. After comparing assembly statistics (total number of transcripts, contig N50, median contig length, and average contig size), version 3 (kmer length 32) was selected because this assembly resulted in the longest N50 (see Table 1). However, assembly version 3 still contained contigs that lacked open reading frames or were expressed at very low levels. Therefore, in order to create an improved assembly that could be used for measuring differential gene expression, this assembly was processed with three additional steps.
First, we used the TransDecoder utility within Trinity [18] to return transcripts that contained an open reading frame (ORF) of at least 100 amino acids. Second, we used the program kallisto [19] to quantify the estimated number counts per transcript and removed any transcripts that had less than 10 counts. Finally, we used BLASTN analyses (E-value cutoff of 10E-20, [20]) to compare the v3 assembly to predicted transcripts in the sugar beet genome (*Beta vulgaris* L., Refbeet v1.2, [21]), representative species of the ten plant clades of Phytozome (www.phytozome.net, version 10), and all sequences available in the GenBank NT database (version 1/22/2016, [22]). Any transcript that was best matched to a non-plant species or had no significant hits was not included in the final assembly. These filtering steps resulted in the final assembly, version 4.

**Differential Expression Analyses**

The individual sample reads were mapped to the version 4 transcriptome assembly using Bowtie [23]. RNA-seq by Expectation-Maximization (RSEM) software [24] was used to account for reads that could re-align to multiple assembled transcripts in the *de novo* assembly due to alternatively spliced isoforms. The raw expression counts were normalized across samples using the Trimmed Mean of M-values (TMM) method [18] in edgeR [25]. GGplot2 (CRAN, [26]) was used to compare and visualize read counts across replicate samples for technical reproducibility. Transcripts in the v4 assembly with a log count per million less than one (log cpm<1) across all samples were excluded from the analyses, leaving 72,697 expressed transcripts (v4 isoforms). edgeR was also used to identify significantly (FDR <0.05) DETs responding to treatment in each genotype at each time point and across all time points.
**Functional Annotation**

The v4 transcriptome was annotated using BLASTX ($E < 10^{-10}$, [20]) against proteins from *Arabidopsis thaliana* (The Arabidopsis Information Resource version 10 [TAIR10], www.arabidopsis.org), sugar beet (Refbeet v1.2, [21]), and Uniref100 (version 1/22/2016, [27]). Custom Perl scripts were used to assign gene ontology (GO) biological processes and molecular function terms [28] based on the top *A. thaliana* hit.

To allow comparisons between the v4 transcriptome assembly and predicted proteins in the grain amaranth and sugar beet genomes, predicted proteins from grain amaranth and sugar beet were also compared to *A. thaliana* (TAIR version 10) using BLASTP ($E < 10^{-10}$ [20]). Custom Perl scripts were then used to assign GO information based on the best *A. thaliana* homolog.

Overrepresented GO terms associated with DETs of interest were identified using a Fisher’s exact test [29] with a Bonferroni correction ($P < 0.05$, [30]) by comparing the number of times each GO term was found within a list DETs of interest relative to the number of times each GO term was found among all transcripts in the v4 assembly.

**Results**

**Phenotypic Assessment of Mesotrione Responses in Resistant and Susceptible Waterhemp Genotypes**

Samples used for RNA-seq were harvested prior to the development of visual mesotrione treatment symptoms; therefore, herbicide-treated and mock-treated control plants were maintained in the greenhouse for three weeks after mesotrione application. Both genotypes responded to the mesotrione application as expected. The resistant population initially displayed the major HPPD-inhibiting herbicide characteristic of
bleached meristematic growth but by the third week after application had recovered. Visual comparison of mock-treated resistant (Figure 1A), mock-treated susceptible (Figure 1B), and mesotrione-treated resistant (Figure 1C) at three weeks after treatment showed slight differences, primarily minor stunting and sparse tissue damage within the canopy. Conversely, the mesotrione-treated susceptible population sustained heavy tissue bleaching and eventually necrosis and plant death (Figure 1D). These observations and comparisons verified the proper herbicide response of both genotypes to mesotrione treatment.

**RNA-Seq and de novo Assembly of the Waterhemp Transcriptome**

Purified RNA from three replicates of 16 samples (3, 6, 12, and 24 hour samples of mesotrione-treated or mock-treated susceptible and resistant genotypes) were sent to the Iowa State University DNA Facility for the creation and sequencing (100 base pair, paired-end sequencing) of 48 multiplex libraries. A total of 2.45 billion raw reads were produced. Following the sequence clean up described in the Materials and Methods, 2.36 billion sequences were used for de novo transcript assembly using the program Trinity (version 2.0.6, [17]) with three different kmer lengths (k = 25, 29, and 32). Sequences from all samples were used to yield a broad representation of the waterhemp transcriptome and allow the identification of genes expressed in a genotype, treatment or time-specific manner.

When comparing the three assemblies (v1, v2 and v3) generated with differing kmer lengths (25, 29, and 32, respectively) we noted that as the kmer length increased, transcript number decreased and Contig N50 increased (Table 1). The Contig N50 is a weighted median of contig (contiguous overlapping sequences) length where 50% of the
assembled nucleotides are contained in contigs greater than or equal to the length of the Contig N50; it can be used as an important measurement in assembly evaluations and was a major factor in the decision of which assembly to use for our analysis [31,32]. In addition, we visualized contig length distribution for each of our different assemblies (Supplemental File 1). As suggested by the contig statistics, increasing kmer size increased average contig length and decreased the number of contigs. This was especially evident for contigs smaller than 1000 base pairs (Log₁₀ 3). Therefore, we chose to focus on the third assembly (v3, kmer = 32) for subsequent analysis. Following selection of the v3 assembly, we still needed to remove sequences that lacked open reading frames (ORFs), were redundant, or were expressed at extremely low levels. From the initial v3 assembly containing 451,199 transcripts, TransDecoder [18] was used to identify all transcriptions with ORFs>100 base pairs and remove redundant transcripts, leaving 128,737 transcripts. Similarly, kallisto [19] identified 97,944 lowly-expressed transcripts in the v3 assembly. Cross referencing the TransDecoder and kallisto datasets resulted in 119,635 transcripts with ORFs>100 bp and read counts >10. BLASTX (E-value <10E-20, [20]) was then used to compare these contigs against proteins in the A. thaliana (L.) Heynh (The Arabidopsis Information Resource version 10 [TAIR10], www.arabidopsis.org), sugar beet (Refbeet v1.2, [21]), and Uniref100 (version 1/22/2016, [27]) databases. Sequences with greatest homology to non-plant species were removed. Following these final BLASTX analyses, we were left with 113,893 transcripts as the basis of our de novo waterhemp transcriptome (v4) used for differential expression analyses. The v4 assembly decreased transcript count and increased Contig N50 to 1,709
bp (Table 1). In addition, many smaller contigs were removed from the assembly (Supplemental File 1).

Annotation of the Waterhemp Transcriptome

The waterhemp v4 assembly was annotated using BLASTX (E<10^{-10}, [20]) against *A. thaliana* (The Arabidopsis Information Resource version 10 [TAIR10], www.arabidopsis.org), sugar beet (Refbeet v1.2, [21]), and Uniref100 (version 1/22/2016, [27]). The best *A. thaliana* hits were used to assign the gene ontology (GO) biological processes and the molecular function terms [28] to each transcript of the v4 assembly. Sequences and annotation for the v4 assembly can be found in supplemental files 2 and 3.

To verify the accuracy and coverage of the v4 assembly, GO biological process terms inferred from homology with *A. thaliana* were mapped to GO slim terms using custom Perl scripts. GO slim term abundance was then compared between the waterhemp v4 transcriptome assembly and all predicted proteins of the *A. thaliana*, sugar beet, and grain amaranth (*Amaranthus hypochondriacus* L.) genomes (Figure 2). Waterhemp, sugar beet and grain amaranth, all belong to the Amaranthaceae family [21,33], while *A. thaliana* is a well-established plant model [34]. For each GO slim term, the abundance of assigned transcripts was measured as a percentage relative to the entire transcriptome or genome, allowing us to normalize for any potential genome duplications within a given species. We found that for thirteen of the fourteen GO slim terms, the v4 waterhemp transcriptome assembly was comparable to the *A. thaliana*, sugar beet and grain amaranth genomes. This suggests the breadth of the waterhemp v4 transcriptome is consistent with the breadth of the *A. thaliana*, sugar beet and grain amaranth genomes. The only
exception was the GO slim term unknown biological process which was overrepresented in \textit{A. thaliana}, compared to the three other species.

**Identification of Waterhemp Transcripts Differentially Expressed in Response to the Mesotrione**

To identify differentially expressed transcripts (DET$s$) responding to mesotrione treatment, individual samples were mapped to the v4 waterhemp assembly using the protocol described in the Trinity user manual (https://github.com/trinityrnaseq/trinityrnaseq/wiki). In total, 782,456,581 reads were mapped to the assembly. The raw expression counts were normalized across samples using the Trimmed Mean of M-values (TMM) method [18] in edgeR [25]. Following visual inspection, all replicate samples were considered good quality. Isoforms were considered expressed if they contained at least 1 count per million across three samples or replicates. Of the 113,893 isoforms in the v4 assembly, 72,697 were considered expressed (v4 isoforms). Average length for the expressed isoforms was 1,580 base pairs and contigs assumed a normal distribution (Supplemental File 1).

edgeR was used to identify DET$s$ responding to mesotrione treatment relative to mock-treated controls within each genotype across time (herbicide resistant and susceptible) and at specific time points (3, 6, 12, and 24 hours after treatment (HAT)). DET expression is reported as a log\textsubscript{2} fold change (log\textsubscript{2} FC). A log\textsubscript{2} FC greater than 1 indicates a DET is induced by the mesotrione treatment, while a log\textsubscript{2} FC less than one indicates a DET is repressed by the mesotrione treatment. DET$s$ with an FDR <0.05 are considered significantly differentially expressed in response to mesotrione treatment (Supplemental File 4).
We identified 89, 62, 61, and 1,983 DETs in the resistant waterhemp genotype at 3, 6, 12, and 24 HAT, respectively, and 500, 77, 61, and 565 DETs were identified in the susceptible waterhemp genotype at 3, 6, 12, and 24 HAT, respectively (Table 2). We plotted the number of DETs per genotype within each time point to analyze expression trends across time (Figure 3). The susceptible waterhemp genotype exhibited large fluxes in DET expression across time. At 3 HAT the susceptible genotype induced 409 transcripts suggesting a quick initial response to the mesotrione treatment. The response diminishes in the middle two time points but then increases again at 24 HAT. In contrast, the resistant waterhemp genotype demonstrated little response to mesotrione treatment at 3, 6 and 12 HAT while a large number of transcripts respond at 24 HAT. Remarkably, while symptoms in response to HPPD herbicide treatments can take as long as one week to develop, both resistant and susceptible waterhemp genotypes responded within three HAT. Furthermore, few DETs overlapped between time points within a given genotype (Figure 4) or between genotypes (Table 2). At 3, 6, 12, and 24 HAT we found 7.7%, 3.7%, 8%, and 3.4% of DETs were common to both waterhemp genotypes, respectively, suggesting a rapid and dynamic response to mesotrione treatment (Table 2).

In addition to identifying transcripts responding to mesotrione treatment at specific time points, we also identified transcripts responding to mesotrione treatment across time. We identified 2,091 and 1,246 DETs responding to mesotrione treatment in the resistant and susceptible genotypes, respectively (Supplemental File 4). Of these, only 330 DETs were common to both waterhemp genotypes. This reaffirms that the resistant and susceptible genotypes have vastly different responses to the mesotrione treatment.
Characterization of Mesotrione Responsive Transcripts

While differential expression is useful in identifying individual transcripts found in response to the mesotrione treatment, we were interested in identifying transcripts responding to mesotrione treatment that might have similar functions or act in the same molecular pathway. Therefore, for each time point by genotype combination, we used a Fisher’s Exact Test [29] with a Bonferroni correction [30] to identify gene ontology biological process terms [35] significantly overrepresented (P<0.05) among DETs, relative to the waterhemp v4 assembly (Supplemental File 5). In the resistant waterhemp genotype, we identified 11 and 12 GO terms significantly overrepresented at 3 and 24 HAT. No significant GO terms were identified at 6 and 12 HAT. Combining all DETs from the resistant waterhemp genotype, we identified 18 significantly overrepresented GO terms. In the susceptible waterhemp genotype, we identified 34, 3, 2 and 24 significant GO terms at 3, 6, 12 and 24 HAT, respectively. Combining all DETs from the susceptible waterhemp genotype, we identified 39 significantly overrepresented GO terms.

To allow direct comparison between resistant and susceptible waterhemp genotypes, we compared unique transcript counts for significant GO terms (P<0.05) identified at specific time points and over time in both genotypes (Figure 5). To aid in data visualization, GO terms with DETs that perfectly overlapped with a larger, significant GO term were removed. In addition, only GO terms with at least 10 DETs in either the resistant or susceptible waterhemp genotype are shown. Using this approach, we were able to identify 18 GO terms significantly overrepresented only in the susceptible waterhemp genotype, nine GO terms significantly overrepresented only in the
resistant waterhemp genotype and nine GO terms significantly overrepresented in both waterhemp genotypes.

GO terms uniquely overrepresented in the susceptible waterhemp genotype response were largely associated with stress and defense responses including responses to osmotic stress (GO:0006970), hyperosmotic salinity (GO: 0042538), other organism (GO:0051707), virus (GO: 0051707), wounding (GO:0009611), and respiratory burst involved in defense response (GO:0002679). Other significantly overrepresented GO terms were associated with metabolism including lignin (GO: 0009809), flavonoid (GO: 0009813), and coumarin (GO:0009805), cellular modified amino acid (GO:0042398), pentacyclic triterpenoid (GO:0019745) and sterol (GO:0016126) biosynthesis, acetyl-CoA (GO: 0006084) and phenylpropanoid (GO:0009698) metabolism, and polyamine catabolism (GO:0006598). Other significant GO terms included protein peptidyl-prolyl isomerization (GO:0000413), peptidyl-proline modification (GO:0018208) and chaperone-mediated protein complex assembly (GO:0051131). For 11 of the 18 significantly overrepresented GO terms unique to the susceptible waterhemp genotype, we observed more DETs in the susceptible than the resistant genotype.

GO terms significantly overrepresented in both waterhemp genotypes included responses to cyclopentenone (GO:0010583), endoplasmic reticulum stress (GO:0034976), hydrogen peroxide (GO:0042542), high light intensity (GO:0009644), reactive oxygen species (GO:0000302), cadmium ion (GO:0046686), and salt stress (GO:0009651) and heat acclimation (GO:0010286)) and toxin catabolism (GO:0009407). For five of the nine GO terms significant in both genotypes, a greater number of DETs was observed in the susceptible waterhemp genotype.
GO terms uniquely overrepresented in the resistant waterhemp genotype were quite varied in their functions. Similar to the responses in the susceptible waterhemp genotype, we identified GO terms associated with response to stress (i.e., hyperosmotic response (GO:0006972) and responses to temperature stimulus (GO:0009266) and karrikin (GO:0080167)). Interestingly, a number of GO terms were associated energy metabolism including amylopectin biosynthesis (GO:0010021), proteasomal protein catabolism (GO:0010498), gluconeogenesis (GO:0006094) and trehalose biosynthesis (GO:0005992). Other significant GO terms observed in the resistant waterhemp genotype included cytoskeleton reorganization (GO:0007010) and transcription (GO:0006351).

To understand how DETs in these GO terms responded to mesotrione treatment, we compared their expression patterns and expression profiles between resistant and susceptible waterhemp genotypes. Of the 4,799 total DETs, 1,311 and 3,034 were uniquely significantly differentially expressed in response to mesotrione treatment in the susceptible waterhemp genotype and the resistant waterhemp genotypes, respectively. A total of 454 DETs were significantly differentially expressed in both genotypes. We then compared DET expression patterns across a core set of nine overrepresented gene ontology terms identified above including cytoskeleton organization, gluconeogenesis, hyperosmotic response, response to cadmium, response to high light intensity, response to salt stress, response to wounding, sterol biosynthesis, and toxin catabolism (Figure 6). When we examined the DETs common to both the resistant and susceptible waterhemp genotypes, we found that the majority of these genes were induced in both genotypes. However, in the susceptible waterhemp genotype, expression was strongly induced 3 HAT, weakly expressed 6 and 12 HAT, and again strongly induced 24 HAT. A similar
response occurred in the resistant waterhemp genotype, however the dip in gene expression observed at 6 and 12 HAT in the susceptible waterhemp genotype was largely restricted to 12 HAT. In contrast, genes repressed in response to mesotrione were weakly repressed at 3, 6 and 12 HAT, but strongly repressed at 24 HAT.

For DETS unique to the susceptible or resistant waterhemp genotypes, we observed differences in the number and expression of DETs depending on the GO terms of interest. The GO terms cytoskeleton organization, gluconeogenesis and trehalose biosynthesis were largely unique to the resistant waterhemp genotype response and were repressed by mesotrione treatment. Few additional DETs, aside from those common to both genotypes, were observed in the susceptible waterhemp genotype. For the GO terms, response to cadmium, salt stress and high light intensity and toxin catabolism, DETs unique to the susceptible waterhemp genotype were largely induced, while DETs associated with these GO terms in the resistant waterhemp genotype were largely repressed. Unique DETs associated with the GO term sterol biosynthesis were repressed in the susceptible waterhemp genotype but had mixed expression in the resistant waterhemp genotype, while unique DETs associated with the GO terms response to wounding and hyperosmotic response had mixed expression among unique susceptible and unique resistant DETs.

Discussion

Weeds are the major pest complex in global agricultural systems and will only continue to be a greater problem with the increasing prevalence of evolved herbicide resistance. Waterhemp is among the worst weeds found in agricultural fields of the Midwestern United States and has populations with evolved resistance to six different
herbicide sites of action including acetolactate synthase (ALS, Herbicide Group (HG) 2, EC 2.2.1.6), 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS synthase, HG 9, EC 2.5.1.19), 4-hydroxyphenylpyruvate dioxygenase (HPPD, HG 27, EC 1.13.11.27), photosystem II (PSII, HG 5, EC 1.10.3.9), protoporphyrinogen oxidase (PPO, HG 14, EC 1.3.3.4) inhibiting herbicides and synthetic auxins (HG 4) [9]. Fields with poorly managed waterhemp can experience yield losses up to 56% in soybean and 74% in maize [2,3]. The ability to evolve herbicide resistance and the potential for yield loss makes this weed an important species to study by the weed science community.

Several *Amaranthus* species, including waterhemp have been the subject of discussion among the weed science community as potential candidates for genomic efforts in weed science [36,37]. Weediness characteristics associated with waterhemp include rapid growth, self-incompatibility, high seed output and dispersal and ability to compete for space and nutrients with crop species. In addition, in 2011, waterhemp populations resistant to five herbicide sites of action were reported in Iowa [38]. Therefore, developing genomic tools to characterize the genes and gene networks involved in herbicide resistance is critically important.

**Development of Genomic Resources for Studying Competitive Traits in Waterhemp**

Initial genomic studies of waterhemp used 454 pyrosequencing to sample the waterhemp genome [15] and develop a waterhemp transcriptome [16]. By sampling less than ten percent of the waterhemp genome, Lee et al. [15] were able to characterize several herbicide resistance target genes. Similarly, Riggins et al. [16] used the 44,469 unique sequences assembled in their transcriptome to identify and characterize a number
of herbicide target site genes for which waterhemp has evolved resistance. However, they pooled RNA samples of different tissues and treatments prior to sequencing, making it impossible to directly differentiate herbicide treatment responses in resistant and susceptible waterhemp genotypes. However, our sequencing and de novo transcriptome assembly approach used 48 multiplexed libraries representing resistant and susceptible waterhemp genotypes, treated and mock-treated with the HPPD herbicide mesotrione across a twenty-four hour time course. Assembling RNA-seq data across libraries allowed us to develop a comprehensive waterhemp leaf transcriptome, which represents an important asset to the weed science community for studying important weedy traits. Assembly statistics of our waterhemp transcriptome (Table 1, Supplemental File 2) coupled with comparisons to predicted proteins from the model species A. thaliana, and the related species sugar beet and grain amaranth (Figure 2) confirm the quality and breadth of our assembly. To increase the utility of the waterhemp transcriptome, supplemental data files include assembled sequences for the v4 assembly and a database of annotated transcripts. Furthermore, raw sequences will be deposited in the National Center for Biotechnology Small Reads Archive (NCBI SRA, http://www.ncbi.nlm.nih.gov/sra) under Bioproject Accession XXXXX. This will allow reassembly and continued improvement as more sequences from waterhemp populations become available.

**Identification of Transcripts Responding to Mesotrione Treatment in Waterhemp**

Mapping of reads from specific waterhemp genotypes, mesotrione and mock treatments and time points allowed us to leverage our waterhemp transcriptome to
identify the genes and gene networks responding to mesotrione treatment in waterhemp.

One of the first trends we observed among DETs was the number of DETs was greatest at 3 and 24 HAT. Furthermore, DETs in both waterhemp genotypes at 3 and 24 HAT were significantly overrepresented with GO terms associated with light (i.e., response to high light intensity, red light and far red light). The samples used in the experiment were collected 12:00 PM CDT, 3:00 PM CDT, 9:00 PM CDT, and 9:00 AM CDT on May 22 and 23 of 2015. Sunrise occurred at 5:49 AM, while sunset occurred at 8:34 PM. The pattern of DETs suggests that light diurnal cycle may play a role in regulating waterhemp responses to light associated herbicides like the HPPD inhibitor herbicides. Since HPPD inhibitor herbicides disrupt photosynthesis and photoprotection in susceptible plant species, the herbicide effect on gene expression would likely decrease when conditions with less available light prevail.

In both the resistant and susceptible waterhemp genotypes, mesotrione treatment resulted in the differential expression of transcripts associated with stress and defense responses. When we compared DET expression in transcripts unique to the resistant or susceptible waterhemp genotypes, we observed that DETs associated with responses to cadmium, high light, hyperosmotic and salt stress and toxin catabolism were largely induced by mesotrione treatment across the time course. In contrast, DETs unique to the resistance waterhemp genotype and associated with these same GO terms were largely repressed across time. These data highlight several remarkable features of resistant waterhemp genotype response to mesotrione. First, responses to mesotrione treatment were detected very quickly, within three HAT in both resistant and susceptible waterhemp genotypes. Second, while the susceptible waterhemp genotype continued to
induce stress responses over the experiment time course, by three HAT the resistant waterhemp genotype was already repressing differential expression of stress-associated genes. This suggested that by three HAT, the resistant waterhemp genotype began to neutralize herbicidal activity and was likely returning to normal physiological function.

While responses to light and stress were expected, our analyses determined that DETs unique to the resistant waterhemp genotype were significantly overrepresented with the GO terms cytoskeleton organization, gluconeogenesis and trehalose biosynthesis. Genes within these GO terms were significantly repressed in response to the mesotrione treatment, especially at 24 HAT. To connect these responses and examine the underlying gene networks, we took advantage of the waterhemp annotation platform to identify the best *A. thaliana* homologs for DETs associated with these GO terms. Unique *A. thaliana* identifiers were then submitted to the String-db website (http://string-db.org) to identify gene networks [39]. Networks in String-db are established using a variety of methods including but not limited to experiments, public databases and co-expression. Collectively, the three GO terms contained 91 DETs which corresponded to 45 unique *A. thaliana* identifiers. Of these, 38 could be assigned to the same network with a high confidence score ranging from 0.71 to 0.99 (Supplemental Files 6). Interestingly, the network also contained several genes associated with herbicide resistance. This included ten aminocyclopropane-1-carboxylic acid (ACC) synthase homologs (AtACS2, AtACS4-AtACS12) and two beta-tubulins inhibitors (AtTUB6 and AtTUB8). Grossman et al. [40] found that silencing an ACC synthase gene in transgenic tomato (*Lycopersicon esculentum* Mill. cv. Hellfrucht/Friihstamm) alleviated the effects of the herbicide quinclorac (3,7-dichloro-8-quinolinecarboxylic acid, HG 4), thus mirroring the
expression we observe in the waterhemp DETs. Dinitroaniline (HG 3) herbicide-resistant tobacco (*Nicotiana tabacum* L.) were generated by co-overexpression of mutant alpha and beta tubulins [41]. These findings suggest that many herbicides may target the same gene networks.

**Conclusion**

Many agronomically important plant species lack well-curated reference genomes. However, RNA-seq has allowed scientific communities to develop transcriptomes for characterizing genes and traits important for agronomic performance of both crops and weeds. In this study, we provide a comprehensive database of the waterhemp transcriptome. While other studies have sequenced the waterhemp transcriptome and identified important herbicide target-sites, this is the first transcriptomic analysis that identifies genes and gene networks that are differentially expressed in response to HPPD inhibiting herbicides in HPPD-resistant and susceptible waterhemp. Our analyses reveal 1) that waterhemp responses to mesotrione are fast and detectable in as little as three hours, 2) resistant and susceptible waterhemp genotypes show little overlap in mesotrione responses and 3) genes targeted by other herbicides may belong to the same gene networks. These findings lay a strong foundation for future research by the weed science community and will improve the opportunities to better manage weeds with evolved resistances to herbicides.

**Acknowledgements**

This research was in part financed by funds provided by the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) CRIS Project 3625-21220-005-
00D, the Iowa Soybean Association (ISA), and the Iowa State University Department of Agronomy. Appreciation is extended to Ms. Lori Lincoln and Dr. Jamie O’Rourke for technical assistance. The USDA is an equal opportunity provider and employer. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U. S. Department of Agriculture.

References


Tables

**Table 1.** Comparison of waterhemp (*Amaranthus tuberculatus*) transcriptome assemblies. Trinity (version 2.0.6, [Grabherr et al. 2011]) was used to generate three different de novo waterhemp transcriptome assemblies: v1, v2, and v3. The three different assemblies differed in the kmer lengths that were required for assembly. A total of 2.3 billion reads representing different genotypes, treatments and time points were used in the assembly. The v4 assembly is a subset of the v3 assembly, with transcripts that were redundant, lacking open reading frames or expressed at low levels were removed. With each assembly, the number of total transcripts decreased while average contig length increased.

<table>
<thead>
<tr>
<th></th>
<th>v1 (k-mer = 25)</th>
<th>v2 (k-mer = 29)</th>
<th>v3 (k-mer = 32)</th>
<th>v4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ‘gene’ count</td>
<td>269,388</td>
<td>238,782</td>
<td>226,402</td>
<td>42,040</td>
</tr>
<tr>
<td>Total transcript count</td>
<td>512,945</td>
<td>471,767</td>
<td>451,199</td>
<td>113,893</td>
</tr>
<tr>
<td><strong>All Transcripts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig N50</td>
<td>762</td>
<td>926</td>
<td>1,029</td>
<td>1,709</td>
</tr>
<tr>
<td>Median Contig Length</td>
<td>390</td>
<td>424</td>
<td>448</td>
<td>1,094</td>
</tr>
<tr>
<td>Average Contig Length</td>
<td>598.54</td>
<td>669.21</td>
<td>713.94</td>
<td>1,317.45</td>
</tr>
<tr>
<td><strong>Longest Isoform</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig N50</td>
<td>662</td>
<td>747</td>
<td>792</td>
<td>1,816</td>
</tr>
<tr>
<td>Median Contig Length</td>
<td>359</td>
<td>371</td>
<td>375</td>
<td>998</td>
</tr>
<tr>
<td>Average Contig Length</td>
<td>549.32</td>
<td>590.09</td>
<td>609.47</td>
<td>1,301.70</td>
</tr>
</tbody>
</table>

**Table 2.** Summary of the differentially expressed transcripts (DETs) responding to mesotrione treatment in waterhemp (*Amaranthus tuberculatus*). DETs were identified at 3, 6, 12 and 24 hours after treatment (HAT) in the resistant and susceptible waterhemp genotypes. Percent overlap between genotypes at a specific time point was calculated by dividing the number of DETs in common between genotypes by the total number of unique DETs at that time point.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Hours After Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Resistant</td>
<td>89</td>
</tr>
<tr>
<td>Susceptible</td>
<td>500</td>
</tr>
<tr>
<td>Overlap</td>
<td>42</td>
</tr>
<tr>
<td>Percent Overlap</td>
<td>7.7%</td>
</tr>
</tbody>
</table>
Figure Legends

**Figure 1.** Phenotyping for mesotrione resistance. To confirm that waterhemp (*Amaranthus tuberculatus*) samples used for RNA-seq analyses were properly treated, a control group of plants was grown simultaneously among the plants sampled for RNA-seq. No tissues were collected from these plants and they were allowed to continue growing for three additional weeks after treatment to assess mesotrione herbicide injury. All plants exhibited the expected phenotype. (A) Resistant genotype, mock-treated (B) Susceptible genotype, mock-treated (C) Resistant genotype, herbicide-treated (D) Susceptible genotype, herbicide-treated.

**Figure 2.** Comparison of the waterhemp (*Amaranthus tuberculatus*) v4 transcriptome assembly to predicted proteins in the *Amaranthus hypochondriacus*, *Beta vulgaris* and *Arabidopsis thaliana* genomes. Gene ontology biological process (GOBP) slim terms were used to compare the breadth of the waterhemp v4 transcriptome to the other species. Within each GOBP slim annotation, the fraction of sequences with that annotation is consistent across all four species, suggesting all four species have relatively equal breadth.

**Figure 3.** Identification of waterhemp (*Amaranthus tuberculatus*) differentially expressed transcripts responding to mesotrione treatment across time. To identify differentially expressed transcripts (DETs) at each time point (3, 6, 12, and 24 hours after treatment (HAT)), transcript expression in resistant (green) or susceptible (magenta) genotypes treated with mesotrione was compared to mock controls. The values above and below the bars represent the number of DETs that were induced and repressed, respectively.

**Figure 4.** Differentially expressed transcripts (DETs) are dynamically expressed in response to mesotrione in waterhemp (*Amaranthus tuberculatus*) across time. (A) Comparison of resistant DETs at 3, 6, 12, and 24 HAT. (B) Comparison of susceptible DETs at 3, 6, 12, and 24 HAT.

**Figure 5.** Characterization of waterhemp (*Amaranthus tuberculatus*) differentially expressed transcripts (DETs) using gene ontology (GO) overrepresentation. A Fisher's exact test (Fisher, 1966) with a Bonferroni correction (Bonferroni, 1935) was used to identify significantly (P<0.05) overrepresented gene ontology biological process terms among DETs relative to all transcripts in the waterhemp v4 transcriptome assembly. GO terms were assigned on the basis of the top *Arabidopsis thaliana* hit (see supplemental file 3). Data was analyzed by each genotype x time point combination and by genotype over all (supplemental file 5). To aid in data visualization, GO terms with DETs that perfectly overlapped with a larger, significant GO term were removed. In addition, only GO terms with at least 10 DETs in either the resistant or susceptible genotype are shown. For each GO term, all significant time points are indicated in parentheses. Black and grey bars indicate DET counts in the resistant and susceptible genotypes, respectively. Data is divided to demonstrate GO processes unique or common to resistant and susceptible genotypes.
Figure 6. Characterizing differentially expressed transcript (DET) expression patterns of waterhemp (*Amaranthus tuberculatus*) in response to mesotrione treatment. A core subset of nine gene ontology (GO) terms identified in Figure 5 were chosen for examining expression of DETs unique to the resistant (unique to R) or susceptible (unique to S) waterhemp genotypes or common to both (common to R and S). Black lines join portions of the heatmaps for particular GO terms of interest. While DETs common to both genotypes are largely induced, DETs unique to resistant or susceptible genotypes tend to have mixed expression patterns. Further, for some GO terms expression patterns are opposite between unique resistant and unique susceptible DETs. Overall, the expression pattern of induced versus repressed DETs is quite different across time.

Supplemental Files

**Supplemental File 1.** Contig length distribution of the four versions of the waterhemp (*Amaranthus tuberculatus*) transcriptome assemblies. Assemblies v1, v2 and v3 were generated with differing kmer lengths (25, 29, and 32, respectively) as described in the materials and methods. Assembly v4 was generated by removing sequences that lacked open reading frames (ORFs), were redundant, or were expressed at extremely low levels in v3. Assembly v4 isoforms contains all transcripts from the v4 assembly with expression counts greater than one count per million in at least three samples or replicates. The v4 isoforms assembly was used for identifying differentially expressed transcripts (DETs). The red line indicates contigs less than 1000 basepair in length. These contigs were the most affect by different assembly parameters.

**Supplemental File 2.** Waterhemp (*Amaranthus tuberculatus*) v4 transcriptome sequences.

**Supplemental File 3.** Annotation of the waterhemp (*Amaranthus tuberculatus*) version 4 transcriptome. BLASTX (Altschul et al., 1997) was used to compare waterhemp transcriptome sequences against the Uniref100 (version 01/22/2016) nonredundant protein database. The generated BLAST report was parsed to identify the top hit and the most informative hit. The most informative hit was identified by eliminating hit descriptions including the words uncharacterized, putative, related, predicted, orf or expressed. Descriptions containing Arabidopsis or Rice gene identifiers (AtXgXXXXX, OsXXgXXXXX) were also ignored. A minimum E-value score E<10⁻¹⁰ was required. The top A. thaliana hit (TAIR version 10) was determined by BLASTX of waterhemp transcriptome sequences against *A. thaliana* proteins (TAIR10, E<10⁻¹⁰). Gene ontology information was inferred from the top *A. thaliana* protein.

**Supplemental File 4.** Identification of waterhemp (*Amaranthus tuberculatus*) transcripts significantly (FDR < 0.05) differentially expressed in response to mesotrione treatment. edgeR was used to identify differentially expressed transcripts (DETs) responding to herbicide treatment relative to mock treated controls with in each genotype (herbicide resistant (R) and susceptible(S)) and at specific time points (3, 6, 12, and 24 hours after...
treatment (HAT)). DET expression is reported as a log2 fold change (log2 FC) with a log2 FC greater than 1 indicating a DET is induced by mesotrione treatment, while a log2 FC less than one indicating a DET is repressed by mesotrione treatment. FDR values highlighted in yellow are significant in the herbicide resistance genotype, while values highlighted in blue are significant in the susceptible genotype. The top *Arabidopsis thaliana* hit is provided for convenience. Full annotation information can be found in Supplemental File 3.

**Supplemental File 5.** Characterization of waterhemp (*Amaranthus tuberculatus*) differentially expressed transcripts (DETs) using gene ontology overrepresentation. A Fisher's exact test (Fisher, 1966) with a Bonferroni correction (Bonferroni, 1935) was used to identify significantly (P<0.05) overrepresented gene ontology biological process terms among DETs relative to all transcripts in the waterhemp v4 transcriptome assembly. GO terms were assigned on the basis of the top *Arabidopsis thaliana* hit (see supplemental file 3). Data was analyzed by each genotype x time point combination and by genotype over all.

**Supplemental File 6.** Identification of gene networks unique to mesotrione resistance associated with the GO terms gluconeogenesis, cytoskeleton organization and trehalose biosynthesis. we took advantage of the waterhemp (*Amaranthus tuberculatus*) annotation platform to identify the best *Arabidopsis thaliana* homologs for 91 unique resistant DETs associated with these GO terms. Forty-five unique *A. thaliana* identifiers were then submitted to the String-db website ([http://string-db.org](http://string-db.org)) to identify potential gene networks. Thirty-eight of the 45 unique *A. thaliana* identifiers were present in the database and could be visualized. Line linking proteins indicate connections between proteins. Different colored lines indicate different data sources.
<table>
<thead>
<tr>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistant</td>
</tr>
<tr>
<td>Mock</td>
</tr>
<tr>
<td>Treatment</td>
</tr>
<tr>
<td>Mesotrione</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>C</td>
</tr>
</tbody>
</table>

Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.
Figure 6.
CHAPTER 4: GENERAL CONCLUSION

Conclusion

Waterhemp (Amaranthus tuberculatus (Moq.) J.D.Sauer) is a major pest throughout the Midwest United States in corn and soybean production. The concern for managing this weed continues to grow as more cases of herbicide resistant populations are documented. To date, waterhemp populations have evolved resistance to six herbicide sites-of-action (Heap 2016). In 2011, two cases of p-hydroxyphenylpyruvate-dioxygenase (HPPD, EC 1.13.11.27) resistance in waterhemp were documented in Iowa and Illinois (Hausman et al. 2011, McMullan and Green 2011). Early studies showed the mechanism of resistance to be increased herbicide metabolism attributable to cytochrome P450 monooxygenase while the inheritance of the resistance was described as polygenic and “complex” (Huffman et al. 2015, Ma et al. 2013). To our knowledge, there are no current studies that examined the transcriptomic expression response of HPPD-herbicide resistance in waterhemp or used a different resistant population and methodology to describe the inheritance of HPPD-resistance.

The objective of the inheritance study was to determine the number of and characterize the allelic expression of the genes responsible for the HPPD-resistance in waterhemp. A known HPPD-resistant waterhemp population was crossed with a known HPPD-susceptible population, bred through two subsequent generations, and all three generations were challenged with an HPPD-inhibitor herbicide (mesotrione) to characterize the trait inheritance. We hypothesized that the resistance trait was polygenic with dominant or semi-dominant allelic characteristics. Our results suggested that the HPPD herbicide resistance in this waterhemp population was polygenic, with more major
loci expressed allowing survival against higher rates of the herbicide. Further, our results also suggested that at least one dominant allele at each active major loci was needed to successfully express the HPPD herbicide resistance trait.

The objective of the transcriptome study was to examine the differences in transcript expression between HPPD-resistant and -susceptible waterhemp populations in response to mesotrione, an HPPD inhibitor herbicide widely used in corn production. Four leaf samples were collected at four time points from HPPD-resistant and -susceptible waterhemp plants treated with mesotrione or water, and RNA-sequencing used to create a de novo transcriptome assembly. Transcript expression between genotypes were compared for the herbicide treatments and time points. We hypothesized that the HPPD herbicide-resistant waterhemp genotype would express different genes than the susceptible genotype or that the genotypes would express similar genes but at different levels of expression. The results demonstrated that the waterhemp response to mesotrione was very rapid, the HPPD-resistant and -susceptible waterhemp genotypes had distinct and different responses to mesotrione and suggested that the possibility existed for overlapping gene networks in response to herbicides with different modes-of-action. Importantly, the RNA sequences and annotations of the waterhemp transcriptome were made publically available for use by other weed scientists.

In conclusion, our complementary studies demonstrated how the response to HPPD-inhibitor herbicides in resistant and susceptible waterhemp populations was rapid and the responses involved a complex interaction of multiple genes. Furthermore, we introduced the idea of overlapping gene networks in response to other herbicide challenges in waterhemp. These studies provided strong evidence of the importance and
effectiveness of incorporating new genetic and genomic approaches to examine herbicide resistance responses in weeds. It is important to have a full understanding how herbicide resistances evolve, function, spread, and most importantly impact agriculture. This knowledge is critical to support better management strategies for the burgeoning global problem of herbicide-resistant weeds. This understanding requires using new and the most current resources available from an array of scientific disciplines to combat the imminent and continuous challenge of herbicide resistance.

References


