

© 2014 Laura A. Chatham

EXAMINATION OF TARGET-SITE-BASED MECHANISMS OF GLYPHOSATE
RESISTANCE IN WATERHEMP (*AMARANTHUS TUBERCULATUS*)

BY

LAURA A. CHATHAM

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Master's Committee:

Professor Patrick J. Tranel, Chair
Professor John A. Juvik
Professor Dean E. Riechers

ABSTRACT

Glyphosate is one of the most important and widely used herbicide in the world. While few weed species had evolved resistance in the two decades after glyphosate's commercialization in 1974, an overreliance on glyphosate and lack of diversity in weed control methods following the introduction of glyphosate-resistant (GR) crops in 1996 has led to increased selection pressure on weeds to evolve resistance. In 2005, the first GR waterhemp (*Amaranthus tuberculatus*) population was identified. Waterhemp is perhaps the most economically threatening weed in the north central United States, with many populations evolving resistance to herbicides spanning multiple sites of action. Three potential mechanisms of resistance have been linked with GR waterhemp: EPSPS gene amplification, EPSPS target-site mutations, and reduced translocation. The objective of this research was to investigate the target-site-based mechanisms (EPSPS gene amplification and EPSPS mutations) of glyphosate resistance. Waterhemp populations were collected throughout Illinois and screened for glyphosate resistance and EPSPS gene amplification to address whether the mechanism could be used as a proxy for the resistance. The majority of the time, resistant populations had EPSPS gene amplification; however, populations without gene amplification were also found. In some populations, an EPSPS target-site mutation conferring a Pro106Ser substitution was associated with resistance, while in others neither of the target-site-based mechanisms was present. Further examination of the association between glyphosate resistance and EPSPS gene amplification was carried out via a multi-state study of GR waterhemp populations. Dose responses were performed at each location and survivors were tested for gene amplification. Four of five populations had EPSPS gene amplification, and one had no amplification but the Pro106Ser substitution instead. In populations with EPSPS gene amplification, copy number appeared to

increase in plants surviving increasing glyphosate rates. Given the incidence of the Pro106Ser mutation in both of these studies, an investigation of this mechanism was carried out to determine the degree to which it confers resistance. A glyphosate dose response carried out on a segregating F₂ population containing the Pro106Ser substitution revealed that a 2- to 4-fold level of resistance is conferred. Gene amplification appears to be the primary mechanism of resistance in the majority of waterhemp populations, with the Pro106Ser substitution present in fewer populations and conferring a lower level of resistance. Results from the studies herein also suggest that additional mechanisms of resistance exist, either separately or in conjunction with the target-site mechanisms studied.

To Mom

ACKNOWLEDGEMENTS

I am pleased to take this opportunity to thank some of the people that have helped me to be successful throughout my time as a Master's student at the University of Illinois. First and foremost, I'd like to thank my advisor, Dr. Pat Tranel for providing me first with the opportunity to succeed and then the guidance to help make it happen, especially when I couldn't see the forest for the trees. I am also grateful to my committee members, Dr. Jack Juvik and Dr. Dean Riechers for their effort in guiding and advising me throughout this process.

To my fellow members of the Tranel lab, past and present, thank you for your friendship and support these past years; I've learned much from all of you. I am especially grateful to Dr. Chance Riggins for his patience and willingness to help me troubleshoot my experiments. To the more recent members of the lab, Janel Huffman, Chenxi Wu, Jiaqi Guo, Jung Eun Song, and Ahmed Sadeque - thank you for your companionship, but also for always being willing to help when asked. I hope I have, or will be able to, return the favor. To Janel ... thank you for always being there to provide friendship and support (and for watering my plants so often on the weekends!). Your enthusiasm and motivation have encouraged and inspired me. To past members of the lab: Edhivia Campos and Alicia Gardner, Edhy ... thank you for first warmly welcoming me into the lab and then for the close friendship that followed, and Alicia ... thank you for your friendship, but also for helping me with even the most tedious of tasks and for providing the conversations that made it all go faster. I'm very appreciative of all the undergraduate workers that have helped out in the past years. I probably sounded like your mother sometimes by asking you so many questions, but I enjoy hearing about your classes, activities, and lives outside the lab while we work! I'd also like to thank Doug Maxwell and Lisa

Gonzini as well as the rest of the weed sciences group, for helping out with my field studies and having the patience to teach me about the applied side of weed science.

Finally, I'd like to thank my parents for their understanding and support throughout my education here at the University of Illinois – be that monetary, emotional, or otherwise. Thank you for always encouraging me to do my best and take the next step. You taught me to always do a little more than is required, to put in a little more effort than is necessary. A few years ago I might not have thanked you for that, but I realize now that my upbringing has contributed largely to my success.

TABLE OF CONTENTS

CHAPTER 1 Introduction.....	1
1.1 Glyphosate	1
1.1.1 Mode of Action	1
1.1.2 Herbicide Characteristics	2
1.1.3 Glyphosate-Resistant Crop Technology	3
1.2 Glyphosate Resistance in Weeds	3
1.2.1 Target-Site-Based Mechanisms of Resistance	4
1.2.2 Non-Target-Site-Based Mechanisms of Resistance.....	6
1.3 Waterhemp.....	7
1.3.1 Waterhemp Biology	7
1.3.2 Herbicide Resistance.....	9
1.4 Glyphosate Resistance in Waterhemp	9
1.5 Research Objectives.....	10
1.6 Attributions	11
1.7 Literature Cited	12
CHAPTER 2 Association of EPSPS Gene Amplification with Glyphosate Resistance in Illinois Waterhemp (<i>Amaranthus tuberculatus</i>) Populations	20
2.1 Abstract.....	20
2.2 Introduction.....	20
2.3 Materials and Methods.....	22
2.3.1 Seed Collection	22
2.3.2 Resistance Testing	23
2.3.3 Examination of Resistance Mechanisms	25
2.4 Results and Discussion	26
2.5 Acknowledgements.....	32
2.6 Literature Cited	32
2.7 Figures.....	36
CHAPTER 3 A Multi-State Study of the Association Between Glyphosate Resistance and EPSPS Gene Amplification in Waterhemp (<i>Amaranthus tuberculatus</i>).....	39

3.1 Abstract.....	39
3.2 Introduction.....	40
3.3 Materials and Methods.....	42
3.3.1 Field Studies.....	42
3.3.2 Examination of Resistance Mechanisms	43
3.3.3 Statistical Analysis.....	45
3.4 Results and Discussion	45
3.4.1 Waterhemp Control.....	45
3.4.2 EPSPS Gene Amplification	45
3.4.3 KY Population	49
3.4.4 Alternative Mechanisms	50
3.5 Acknowledgements.....	51
3.6 Literature Cited	52
3.7 Table and Figures.....	56
CHAPTER 4 Examination of the EPSPS Pro106Ser Amino Acid Substitution in	
Glyphosate-Resistant Waterhemp (<i>Amaranthus tuberculatus</i>)	62
4.1 Abstract.....	62
4.2 Introduction.....	63
4.3 Materials and Methods.....	65
4.3.1 Creation of a Segregating Line	65
4.3.2 Screening for Pro106Ser Substitution.....	66
4.3.3 Glyphosate Dose Response.....	66
4.3.4 Dose Response Statistical Analysis	69
4.4 Results and Discussion	69
4.5 Literature Cited	73
4.6 Table and Figures.....	77
CHAPTER 5 Concluding Remarks	
5.1 Summary and Conclusions	80
5.2 Implications and Future Directions.....	82
5.3 Literature Cited	87

APPENDIX A A User's Guide to High-Throughput Screening for Herbicide

Resistance in Waterhemp.....	91
A.1 Cone-tainers	91
A.2 Soil	92
A.3 Filling Cone-tainers.....	93
A.4 Germination	94
A.5 Transplanting	95
A.6 Plant Care.....	96
A.6.1 Fertilizing.....	96
A.6.2 Overhead Misting.....	97
A.7 Spraying	98
A.8 Rating.....	99
A.9 Disposal and Cleanup.....	99
A.10 Out-of-Season Experiments	100
A.11 Literature Cited	100
A.12 Tables and Figures	101
APPENDIX B Annotated R Code Used for Statistical Analysis and Creation of Figures.....	114
B.1 Introduction	114
B.2 Basics	114
B.2.1 Basic Code.....	115
B.3 Statistical Analysis and Figures Presented in Chapter 2	117
B.3.1 Code For Determining a Threshold Copy Number.....	117
B.3.2 Map of Illinois Code	120
B.3.3 Proportion Bar Graph Code.....	123
B.4 Statistical Analysis and Figures Presented in Chapter 3	127
B.4.1 Multi-State Data Analysis and Bootstrapping.....	127
B.4.2 Map of Study Sites Code.....	134
B.4.3 Percent Control Ratings Plot Code.....	137
B.4.4 Code For Figure 3.4 –Boxplots, Interaction Plots, and Violin Plots.....	140
B.4.5 Stacked Bar Graph of Pro106Ser Percentages	149

B.5 Statistical Analysis and Figures Presented in Chapter 4	152
B.5.1 Analysis of Dose Response Data and Creation of Dose Response Curves.....	152
B.6 Literature Cited	163

CHAPTER 1

Introduction

1.1 Glyphosate

Glyphosate (N-(phosphonomethyl)glycine) was first discovered at a Swiss pharmaceutical company by Dr. Henri Martin, but having no pharmaceutical use was never reported. Dr. John Franz of Monsanto Co. was the first to synthesize and test the compound as an herbicide in 1970. Glyphosate, as its isopropylamine salt, was first commercialized by Monsanto Co. as Roundup in 1974 (Duke and Powles 2008). It quickly grew in popularity and has since been considered one of the most ideal herbicides for a multitude of reasons.

1.1.1 Mode of Action

The shikimic acid pathway, present in bacteria, fungi and the chloroplasts of plants, has been described as a “metabolic tree with many branches”, producing an extensive number of essential biochemical compounds including vitamins E and K, folic acid, ubiquinone, plastoquinone, the aromatic amino acids tryptophan, phenylalanine, and tyrosine, as well as many secondary metabolites (Bentley and Haslam 1990). Glyphosate inhibits EPSPS, the enzyme catalyzing the formation of 5-enolpyruvylshikimate-3-phosphate (EPSP) from shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP), the sixth step in the shikimate pathway (Steinrücken and Amrhein 1980). Glyphosate acts as a competitive inhibitor of EPSPS with respect to PEP and uncompetitively with respect to S3P (Steinrücken and Amrhein 1980, Boocock and Coggins 1983), creating an enzyme : substrate : inhibitor complex, with glyphosate

in place of PEP (Gruys et al. 1992). Blocking EPSPS activity prevents the synthesis of all downstream products and results in reduced feedback inhibition, causing a massive carbon flux into the shikimic acid pathway and leading to a buildup of shikimate (Duke and Powles 2008). While the exact method by which glyphosate leads to plant death remains somewhat unclear, most believe that a lack of aromatic amino acids preventing protein synthesis starves the plant to death. However, there is also evidence for plant death resulting from the massive carbon flux away from other essential pathways due to the deregulation of the shikimate pathway (Servaites 1987, Siehl 1997, Duke and Powles 2008).

1.1.2 Herbicide Characteristics

One of the attributes that contributes largely to glyphosate's effectiveness in killing plants is its excellent uptake and translocation. Glyphosate is quickly absorbed at plant surfaces (Caseley and Coupland 1985, Kirkwood et al. 2000) and after entering the sieve-tube element, travels through the phloem from source to sink, arriving at meristematic tissues in both roots and shoots (Sprankle et al. 1973, Dewey 1981, Gougler and Geiger 1981). The slow mode of action in addition to the systemic nature of glyphosate makes it an especially efficient herbicide (Duke and Powles 2008), capable of killing plants that contact herbicides often miss, such as perennials and mature plants.

Because the glyphosate target site is only present in some bacteria, fungi, and in plants, glyphosate has very low mammalian toxicity. It also has an excellent environmental profile, with little movement through the soil to groundwater, and a fairly short half-life in the soil as a result of microbial degradation (Duke and Powles 2008). In addition, glyphosate has a broad spectrum of activity; there have been no reports of any plants being naturally resistant to glyphosate,

although some are more tolerant than others. From an economic standpoint glyphosate is also ideal. For producers it is fairly inexpensive to synthesize. For consumers it provides excellent weed control at a very affordable price now that the patent has expired and generic forms are readily available.

1.1.3 Glyphosate-Resistant Crop Technology

Glyphosate's largest shortcoming was its non-selectivity. For the first two decades after its commercialization glyphosate was primarily used as a burndown herbicide to control weeds prior to planting or in situations where plant foliage could be avoided, as in orchards or vineyards (Duke and Powles 2008). Glyphosate-resistant (GR) crop technology was first commercialized in 1996 in soybeans and was accomplished using an insensitive EPSPS enzyme encoded by the CP4 gene isolated from *Agrobacterium* sp. The CP4-EPSPS has a high tolerance for glyphosate while maintaining a high affinity for PEP, allowing the shikimic acid pathway to proceed normally in plants with the CP4 insertion. The CP4 gene is now used in the majority of GR crops (Dill 2005). Within a decade of their release, 96% of all soybean and 61% of all maize planted in the United States were glyphosate resistant (Dill et al. 2008).

1.2 Glyphosate Resistance in Weeds

Prior to the advent of GR crop technology, evolved glyphosate resistance in a weed species was thought to be highly unlikely since there were no cases of resistance yet, there was no evidence for glyphosate metabolism in plants, and it was thought that any mutation in EPSPS would likely result in such a heavy fitness penalty that resistance would never get a foothold

(Bradshaw et al. 1997). However, the first case of evolved glyphosate resistance was identified shortly after the introduction of GR crops in a population of ryegrass (*Lolium rigidum*) from an orchard in New South Wales, Australia at which multiple glyphosate applications had been made consecutively for 15 years (Powles et al. 1998). With the commercialization of GR crop technology released in 1996 and the expiration of glyphosate's patent in 2000, the use of glyphosate increased significantly, resulting in increased selection pressure on weeds to evolve resistance. Horseweed (*Conyza canadensis*) in 2000 was the first weed to evolve resistance in a GR crop (VanGessel 2001). Twenty-eight different weed species across 24 countries have now evolved resistance to glyphosate (Heap 2014).

1.2.1 Target-Site-Based Mechanisms of Resistance

A mutation in the EPSPS gene resulting in a proline to serine amino acid substitution at position 106 (P106S) was first found to be responsible for glyphosate resistance in Malaysian goosegrass (*Eleusine indica*) populations (Baerson et al. 2002). A threonine substitution for the same proline was also found to confer resistance in these goosegrass populations (Ng et al. 2003). Subsequently, proline to threonine (P106T), proline to alanine (P106A), proline to serine (P106S), and proline to leucine (P106L) substitutions were discovered in populations of GR ryegrass (*Lolium rigidum*) (Wakelin and Preston 2006; Yu et al. 2007, Simarmata and Penner 2008, Kaundun et al. 2011). Proline to serine and proline to alanine substitutions have now been found in Italian ryegrass (*Lolium multiflorum*) (Perez-Jones et al. 2007, Jasieniuk et al. 2008). P106S mutations were recently identified in waterhemp (*Amaranthus tuberculatus*) (Bell et al. 2013, Nandula et al. 2013), and junglerice (*Echinochloa colona*) (Alarcón-Reverte et al. 2013). Threonine substitutions have also recently been identified in sourgrass (*Digitaria insularis*) (de

Carvalho et al. 2012) and Sumatran fleabane (*Conyza sumatrensis*) (González-Torralva et al. 2014). Further research into this mechanism of resistance has indicated that small amino acid substitutions at the analogous proline in *E. coli* result in a narrowing of the glyphosate-binding site of EPSPS, allowing the enzyme to maintain its catalytic efficiency with the substrate PEP, but discouraging binding with the larger inhibitor molecule, glyphosate (Healy-Fried et al. 2007). Questions regarding the potential fitness penalty of these proline 106 substitutions in higher plants have not yet been resolved, but many assume that mutations lowering the affinity of glyphosate for EPSPS will also decrease the affinity of PEP for EPSPS.

An alternative mechanism of target-site-based glyphosate resistance was discovered in Palmer amaranth (*Amaranthus palmeri*) to be amplification of the EPSPS gene; resistant plants had between 5- and 160-fold more copies compared to sensitive plants (Gaines et al. 2010). Genomic EPSPS copy number was shown to be positively correlated with EPSPS transcript level, protein expression and protein activity (Gaines et al. 2010, Ribeiro et al. 2014). The production of extra functional EPSPS proteins in the resistant plant creates additional EPSPS target sites for both the substrate and inhibitor. In this way, the extra EPSPS enzymes act as a “molecular sponge”, soaking up glyphosate and allowing the shikimic acid pathway to proceed normally (Powles 2010). The mechanism of amplification in Palmer amaranth remains unclear, however research suggests the mechanism is DNA mediated (Gaines et al. 2013). EPSPS gene amplification has since been reported in waterhemp (Tranel et al. 2010), Italian ryegrass (Salas et al. 2012), kochia (*Kochia scoparia*) (Wiersma 2012), and spiny amaranth (*Amaranthus spinosus*) (Nandula et al. 2014). There is the potential for a fitness penalty among species with EPSPS gene amplification – the synthesis of additional EPSPS copies may alter carbon flux through other metabolic pathways, and the potential for transposon mediated amplification of the gene may

result in the disruption of other vital genes. While studies regarding the fitness of plants with EPSPS gene amplification have not yet been carried out in all species with this mechanism of resistance, studies in Palmer amaranth found no evidence of a fitness penalty in plants with elevated EPSPS copy number (Giacomini et al. 2014, Vila-Aiub et al. 2014)

1.2.2 Non-Target-Site-Based Mechanisms of Resistance

The first case of altered glyphosate translocation was found in ryegrass; instead of accumulating in the roots as in sensitive plants, glyphosate accumulated in the leaf tips of resistant plants (Lorraine-Colwill 2002). Altered translocation was also discovered in GR horseweed populations (Koger and Reddy 2005), and further examination of the mechanism of resistance was carried out using ^{31}P NMR techniques, which revealed rapid vacuolar sequestration, preventing glyphosate entry into the phloem and therefore translocation to sink tissues throughout the plant (Ge et al. 2010). Vacuolar sequestration of glyphosate has now been confirmed in *Lolium* spp. as well (Ge et al. 2012). Restricted uptake has subsequently been suggested as a mechanism of glyphosate resistance in Palmer amaranth, waterhemp and johnsongrass (*Sorghum halepense*) (Vila-Aiub et al. 2012, Sammons and Gaines 2014). More recently, restricted entry into the chloroplast was hypothesized as the mechanism of glyphosate resistance in *Desmanthus illinoensis*; this mechanism has been described as perhaps the most potent of the glyphosate resistance mechanisms known to date (Sammons and Gaines 2014). An unusual type of non-target-site resistance has been observed in populations of giant ragweed (*Ambrosia trifida*). These plants treated with glyphosate show a hypersensitive response similar to that observed after pathogen attack, with rapid necrosis of herbicide-treated leaves and

resulting in reduced translocation to sink tissues, which allows regrowth from axillary meristems (Robertson 2010, Segobye 2013).

1.3 Waterhemp

Waterhemp (*Amaranthus tuberculatus*) is a dioecious weed species that, although native to Illinois (Sauer 1955), has only been a major pest for the past quarter of a century (Steckel 2007). Waterhemp is highly competitive with crops grown in the Midwest. It can reduce soybean yields up to 43% (Hager et al. 2002) and corn yields up to 74% with season long competition (Steckel and Sprague 2004). While chemical control has been able to keep most populations at bay in the past, the evolution of herbicide resistance and the potential for stacked resistances in waterhemp threatens the use of herbicides for effective waterhemp control.

1.3.1 Waterhemp Biology

A number of characteristics contribute to waterhemp's success in competing with crops. Waterhemp is a summer annual, completing its entire life cycle at the same time as the crops grown in the Midwest allowing it to be highly competitive with crops. Emergence is prolonged throughout the growing season, with seedlings emerging in flushes from May to August (Hartzler et al. 1999). Waterhemp also uses the C4 photosynthetic pathway and therefore has a higher photosynthetic rate in situations with high light and high temperature and it has better water use efficiency. It has an indeterminate growth pattern and plant morphology may vary based on environmental conditions; however, when conditions are ideal, shoots may reach up to 3 m in length (Costea et al. 2005)

Waterhemp is dioecious, having separate male and female plants. Male plants have flowers lacking gynoecium and only produce pollen, while female plants have flowers lacking androecium and are only capable of producing seed when pollinated by a male plant. This obligate outcrossing increases genetic diversity in the species, but may prevent the colonization of new areas if both sexes are not present, since isolated female plants are incapable of producing seed (Tranel and Trucco 2009). This disadvantage could be mediated by the fact that waterhemp is very efficiently wind pollinated with highly aerodynamic pollen grains that may travel more than 300 m from a male plant in ideal conditions (Costea et al. 2005).

After successful pollination, a single female waterhemp plant may produce up to one million seeds (Steckel et al. 2003). Seeds are very small and lightweight and are easily dispersed by water, birds, animals and anthropomorphic activities like the movement of farm machinery or the use of manure and compost (Costea et al. 2005). Seed dormancy allows waterhemp seeds to form a persistent seed bank. One study found seeds germinating as many as 17 years after burial (Burnside et al. 1996) and others reported that 11% of seeds retain viability after four years of burial close to the soil surface (Buhler and Hartzler 2001).

The above characteristics in combination with management styles adopted in the past few decades may have contributed to waterhemp's invasion and spread in Midwest fields. While tillage would bury small seeds and prevent germination, the reduced tillage practices adopted in recent years have favored small-seeded species like waterhemp that germinate better when closer to the soil surface. More reliance on postemergence instead of preemergence herbicides may also have favored waterhemp's invasion. Postemergence herbicides without soil residual activity are only capable of controlling weeds present at the time of spraying. However, waterhemp's

prolonged emergence pattern makes it a difficult weed to control with a single postemergence application (Costea et al. 2005, Tranel and Trucco 2009).

1.3.2 Herbicide Resistance

Waterhemp has a high propensity to evolve resistance and to date has evolved resistance to herbicides targeting six different sites of action including: acetolactate synthase (ALS), protoporphyrinogen oxidase (PPO), photosystem II (PSII), EPSPS, 4-hydroxyphenylpyruvate dioxygenase (HPPD) and, most recently, plant growth regulators (Heap 2014). Waterhemp populations with quad-stack and quint-stack resistance have recently been identified in Illinois and Missouri (Bell et al. 2013, Schultz et al. 2014, Heap 2014).

1.4 Glyphosate Resistance in Waterhemp

The first reported cases of GR waterhemp were found in Missouri fields that had been planted with continuous GR soybean for at least six years with often more than one yearly glyphosate application (Legleiter and Bradley 2008). Since 2005, GR waterhemp has been documented in 15 states.

Several mechanisms of glyphosate resistance in waterhemp have been proposed in recent years. Gene amplification of the target-site gene EPSPS was observed in Missouri waterhemp populations by Bell et al. (2009). While believed to be associated with resistance, the extent to which elevated EPSPS copy number contributes to glyphosate resistance remains unknown (Bell et al. 2009, Tranel et al. 2010, Shaner et al. 2012). Bell et al. (2009) found plants with elevated

EPSPS copy number that were not resistant, suggesting that EPSPS gene amplification alone may not be sufficient to confer resistance.

A single nucleotide polymorphism replacing a thymine for cytosine in the EPSPS gene, resulting in a proline to serine amino acid substitution has also been identified in GR waterhemp populations (Bell et al. 2013, Nandula et al. 2013). However, questions remain regarding the relationship between the point mutation and glyphosate resistance in waterhemp. Nandula et al. (2013) found evidence for non-target-site-based mechanisms working in addition to the P106S substitution in a GR waterhemp population from Mississippi. These non-target-site mechanisms included a reduced absorption pattern 24 hours after treatment, and significantly lower glyphosate translocation out of treated leaves.

1.5 Research Objectives

While preliminary reports recognize the presence of EPSPS gene amplification and a target-site mutation resulting in proline to serine substitution in populations of GR waterhemp, the extent to which they confer resistance remains somewhat unclear. The overarching goal of this research was to gain a better understanding of target-site-based mechanisms of glyphosate resistance in waterhemp.

The testing and confirmation of GR waterhemp populations is a valuable tool that may allow farmers to make more informed decisions regarding the management of their fields. However, whole-plant screening of waterhemp populations in the greenhouse is inefficient and rather tedious. The objective of the research described in Chapter 2 was, given the potentially uncertain relationship between EPSPS copy number and resistance level, to determine whether EPSPS gene amplification can be used as a proxy for glyphosate resistance. Populations

throughout the state of Illinois were collected, screened in the greenhouse for glyphosate resistance, and survivors were tested for EPSPS gene amplification.

To expand on previous studies of EPSPS gene amplification in waterhemp, the objective of the research described in Chapter 3 was to determine whether gene amplification is in fact associated with resistance and further to examine the relationship between copy number and glyphosate rate. Glyphosate dose response studies were carried out at sites with GR waterhemp. Survivors were sampled and tested for EPSPS gene amplification. A secondary objective of the research conducted in both Chapters 2 and 3 was to document the distribution of GR waterhemp populations as well as the mechanisms responsible for resistance.

The objective of the research in Chapter 4 was to further document the relationship between glyphosate resistance and the Pro106Ser substitution observed in waterhemp. A segregating population was created and dose response experiments were carried out on wild type, heterozygous, and homozygous plants with the target-site mutation in order to document the level of resistance conferred.

1.6 Attributions

The waterhemp populations described in Chapter 2 were collected by Bryan Young, Gordon Roskamp, Patrick Tranel and Aaron Hager. Greenhouse screening of waterhemp populations was carried out in collaboration with Chenxi Wu. The primers used in qPCR for amplifying *EPSPS* were originally designed by Gaines et al. (2010) for use in Palmer amaranth. The primers used for amplification of the reference gene (encoding the large subunit of carbamoyl phosphate synthetase) were designed by Chance Riggins, who also optimized the qPCR assay for waterhemp used in Chapters 2 and 3. The Pro106Ser dCAPS assay described in

Chapters 2, 3, and 4 was designed by Chance Riggins as well. The material presented in Chapter 2 was submitted for publication in Weed Technology in collaboration with Chenxi Wu, Chance Riggins, Aaron Hager, Bryan Young, Gordon Roskamp and Patrick Tranel.

The dose-response field studies conducted outside Illinois described in chapter 3 were performed by Kevin Bradley and colleagues from the University of Missouri, Greg Kruger and colleagues from the University of Nebraska-Lincoln, James Martin and colleagues from the University of Kentucky, and Dallas Peterson, Mithila Jugulam and colleagues from Kansas State University. Tissue samples were shipped to the University of Illinois, where I processed the samples. The Illinois dose responses from 2012 and 2013 were carried out with help from weed science staff at the University of Illinois, including Doug Maxwell and Lisa Gonzini. Mike Owen and colleagues at the University of Iowa completed a dose response in Iowa in 2012 which was used in part to design the 2013 study. Material presented in Chapter 3 will be submitted for publication in Weed Science in collaboration with Kevin Bradley, Greg Kruger, James Martin, Mike Owen, Dallas Peterson, Mithila Jugulam and Patrick Tranel.

Crosses for the pseudo-F₂ population described in Chapter 4 were made by Jiaqi Guo, and the F₁ population used in making the F₂s was made by Sukhvinder Singh.

1.7 Literature Cited

Alarcón-Reverte R, García A, Urzúa J, Fischer AJ (2013) Resistance to glyphosate in junglerice

(*Echinochloa colona*) from California. Weed Sci 61: 48-54

Baerson SR, Rodriguez DJ, Tran M, Feng Y, Biest NA, Dill GM (2002) Glyphosate-resistant

goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. Plant Physiol 129: 1265-1275

- Bell MS, Hager AG, Tranel PJ (2013) Multiple resistance to herbicides from four site-of-action groups in waterhemp (*Amaranthus tuberculatus*). *Weed Sci* 61: 460-468
- Bell MS, Tranel PJ, Riggins CW (2009) Glyphosate resistance in waterhemp: inheritance and EPSPS copy number. *North Central Weed Sci Soc Proc* 64:100
- Bentley R, Haslam E (1990) The shikimate pathway-a metabolic tree with many branches. *Crit Rev Biochem Mol Biol* 25: 307-384
- Boocock MR, Coggins JR (1983) Kinetics of 5-enolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate. *FEBS Lett* 154: 127-133
- Bradshaw LD, Padgett SR, Kimball SL, Wells BH (1997) Perspectives on glyphosate resistance. *Weed Technol* 189-198
- Buhler DD, Hartzler RG (2001) Emergence and persistence of seed of velvetleaf, common waterhemp, woolly cupgrass, and giant foxtail. *Weed Sci* 49: 230-235
- Burnside OC, Wilson RG, Weisberg S, Hubbard KG (1996) Seed longevity of 41 weed species buried 17 years in eastern and western Nebraska. *Weed Sci* 74-86
- Caseley J, Coupland D (1985) Environmental and plant factors affecting glyphosate uptake, movement and activity. In: Grossbard, E., Atkinson, D. (Eds.), *The Herbicide Glyphosate*. Butterworths, London, pp. 92-124.
- Costea M, Weaver SE, Tardif FJ (2005) The biology of invasive alien plants in Canada. 3. *Amaranthus tuberculatus* (Moq.) Sauer var. *rudis* (Sauer) Costea & Tardif. *Can J Plant Sci* 85: 507-522
- de Carvalho LB, Alves PLCA, González-Torralva F, Cruz-Hipolito HE, Rojano-Delgado AM, De Prado R, Gil-Humanes J, Barro F, Luque de Castro MD (2012) Pool of resistance mechanisms to glyphosate in *Digitaria insularis*. *J Agric Food Chem* 60: 615-622

- Dewey SA (1982) Manipulation of assimilate transport patterns as a method of studying glyphosate translocation in tall morningglory [*Ipomoea purpurea* (L.) Roth]. Dissertation Abstracts International, B 42(5): 1695-1696
- Dill GM (2005) Glyphosate-resistant crops: history, status and future. *Pest Manag Sci* 61: 219-224
- Dill GM, CaJacob CA, Padgett SR (2008) Glyphosate-resistant crops: adoption, use and future considerations. *Pest Manag Sci* 64: 326-331
- Duke SO, Powles SB (2008) Glyphosate: a once-in-a-century herbicide. *Pest Manag Sci* 64: 319-325
- Gaines TA, Wright AA, Molin WT, Lorentz L, Riggins CW, Tranel PJ, Beffa R, Westra P, Powles SB (2013) Identification of genetic elements associated with EPSPS gene amplification. *PLoS one* 8: e65819
- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc Natl Acad Sci U.S.A.* 107:1029-1034
- Ge X, d'Avignon DA, Ackerman JJ, Collavo A, Sattin M, Ostrander EL, Hall EL, Sammons RD, Preston C (2012) Vacuolar glyphosate-sequestration correlates with glyphosate resistance in ryegrass (*Lolium* spp.) from Australia, South America, and Europe: a ³¹P NMR investigation. *J Agric Food Chem* 60: 1243-1250
- Ge X, d'Avignon DA, Ackerman JJH, Sammons RD (2010) Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. *Pest Manag Sci* 66:345–348

- Giacomini D, Westra P, Ward SM (2014) Impact of genetic background in fitness cost studies: an example from glyphosate-resistant Palmer amaranth. *Weed Sci* 62: 29-37
- González-Torralva F, Gil-Humanes J, Barro F, Domínguez-Valenzuela JA, De Prado R (2014) First evidence for a target site mutation in the EPSPS2 gene in glyphosate-resistant Sumatran fleabane from citrus orchards. *Agron Sustain Dev* 34:553-560
- Gougler JA, Geiger DR (1981) Uptake and distribution of N-phosphonomethylglycine in sugar beet plants. *Plant Physiol* 68: 668-672
- Gruys KJ, Walker MC, Sikorski JA (1992) Substrate synergism and the steady-state kinetic reaction mechanism for EPSP synthase from *Escherichia coli*. *Biochemistry* 31:5534-5544
- Hager AG, Wax LM, Stoller EW, Bollero GA (2002) Common waterhemp (*Amaranthus rudis*) interference in soybean. *Weed Sci* 50:607-610
- Hartzler RG, Buhler DD, Stoltenberg DE (1999) Emergence characteristics of four annual weed species. *Weed Sci* 578-584
- Healy-Fried ML, Funke T, Priestman MA, Han H, Schonbrunn E (2007) Structural basis of glyphosate tolerance resulting from mutations of Pro101 in *Escherichia coli* 5-enolpyruvylshikimate-3-phosphate synthase. *J Biol Chem* 282: 32949-32955
- Heap IM (2014) International Survey of Herbicide Resistant Weeds. www.weedscience.org. Accessed June 17, 2014
- Jasieniuk M, Ahmad R, Sherwood AM, Firestone JL, Perez-Jones A, Lanini WT, Mallory-Smith C, Stednick Z (2008) Glyphosate-resistant Italian ryegrass (*Lolium multiflorum*) in California: distribution, response to glyphosate, and molecular evidence for an altered target enzyme. *Weed Sci* 56: 496-502

- Kaundun SS, Dale RP, Zelaya IA, Dinelli G, Marotti I, McIndoe E, Cairns A (2011) A novel P106L mutation in EPSPS and an unknown mechanism (s) act additively to confer resistance to glyphosate in a South African *Lolium rigidum* population. J Agric Food Chem 59: 3227-3233
- Kirkwood RC, Hetherington R, Reynolds TL, Marshall G (2000) Absorption, localisation, translocation and activity of glyphosate in barnyardgrass (*Echinochloa crus-galli* (L) Beauv): influence of herbicide and surfactant concentration. Pest Manag Sci 56: 359-367
- Koger CH, Reddy KN (2005) Role of absorption and translocation in the mechanism of glyphosate resistance in horseweed (*Conyza canadensis*). Weed Sci 53: 84-89
- Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. Weed Sci 56: 582-587
- Lorraine-Colwill D, Powles S, Hawkes TR, Hollinshead P, Warner S, Preston C (2002) Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. Pestic Biochem Physiol 74: 62-72
- Nandula VK, Ray JD, Ribeiro DN, Pan Z, Reddy KN (2013) Glyphosate resistance in tall waterhemp (*Amaranthus tuberculatus*) from Mississippi is due to both altered target-site and nontarget-site mechanisms. Weed Sci 61: 374-383
- Nandula VK, Wright AA, Bond JA, Ray JD, Eubank TW, Molin WT (2014) EPSPS amplification in glyphosate-resistant spiny amaranth (*Amaranthus spinosus*): a case of gene transfer via interspecific hybridization from glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*). Pest Manag Sci 10.1002/ps.3754

- Ng C, Wickneswari R, Salmijah S, Teng Y, Ismail B (2003) Gene polymorphisms in glyphosate-resistant and-susceptible biotypes of *Eleusine indica* from Malaysia. *Weed Res* 43: 108-115
- Perez-Jones A, Park K, Polge N, Colquhoun J, Mallory-Smith CA (2007) Investigating the mechanisms of glyphosate resistance in *Lolium multiflorum*. *Planta* 226: 395-404
- Powles SB (2010) Gene amplification delivers glyphosate-resistant weed evolution. *Proc Natl Acad Sci USA* 107: 955-956
- Powles SB, Lorraine-Colwill DF, Dellow JJ, Preston C (1998) Evolved resistance to glyphosate in rigid ryegrass (*Lolium rigidum*) in Australia. *Weed Sci* 604-607
- Ribeiro DN, Pan Z, Duke SO, Nandula VK, Baldwin BS, Shaw DR, Dayan FE (2014) Involvement of facultative apomixis in inheritance of EPSPS gene amplification in glyphosate-resistant *Amaranthus palmeri*. *Planta* 239: 199-212
- Robertson RR (2010) Physiological and biochemical characterization of glyphosate resistant *Ambrosia trifida* L. Master's Thesis. West Lafayette, IN: Purdue University. 87p
- Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW, Scott RC, Burgos NR (2012) EPSPS gene amplification in glyphosate-resistant Italian ryegrass (*Lolium perenne* ssp. *multiflorum*) from Arkansas. *Pest Manag Sci* 68: 1223-1230
- Sammons RD, Gaines TA (2014) Glyphosate resistance: state of knowledge. *Pest Manag Sci* 10.1002/ps.3743
- Sauer JD (1955) Revision of the dioecious amaranths. *Madroño* 13: 5-46
- Schultz J, Riley EB, Wait JD, Bradley KW (2014) Distribution of multiple herbicide resistance in Missouri waterhemp populations. *Weed Sci Soc Am Proc* 54:290

- Segobye K (2013) Biology and ecology of glyphosate-resistant giant ragweed (*Ambrosia trifida* L.) Master's Thesis. West Lafayette, IN: Purdue University. 169p
- Servaites JC, Tucci MA, Geiger DR (1987) Glyphosate effects on carbon assimilation, ribulose biphosphate carboxylase activity, and metabolite levels in sugar beet leaves. *Plant Physiol* 85: 370-374
- Shaner DL, Lindenmeyer RB, Ostlie MH (2012) What have the mechanisms of resistance to glyphosate taught us? *Pest Manag Sci* 68:3-9
- Siehl D (1997) Inhibitors of EPSP synthase, glutamine synthetase and histidine synthesis. *Rev Toxicol* 1: 37-68
- Simarmata M, Penner D (2008) The basis for glyphosate resistance in rigid ryegrass (*Lolium rigidum*) from California. *Weed Sci* 56: 181-188
- Sprankle P, Penner D, Meggitt W (1973) The movement of glyphosate and bentazon in corn, soybean, and several weed species. *Abstr Weed Sci Soc Amer* 13: 75
- Steckel LE (2007) The dioecious *Amaranthus* spp.: here to stay. *Weed Technol* 21: 567-570
- Steckel LE, Sprague CL (2004) Common waterhemp (*Amaranthus rudis*) interference in corn. *Weed Sci* 52:359-364
- Steckel LE, Sprague CL, Hager AG, Simmons FW, Bollero GA (2003) Effects of shading on common waterhemp (*Amaranthus rudis*) growth and development. *Weed Sci* 51:898-903
- Steinrücken H, Amrhein N (1980) The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid-3-phosphate synthase. *Biochem Biophys Res Commun* 94: 1207-1212
- Tranel PJ, Riggins CW, Bell MS, Hager AG (2010) Herbicide resistances in *Amaranthus tuberculatus*: a call for new options. *J Agric Food Chem* 59: 5808-5812

- Tranel PJ, Trucco F (2009) 21st-century weed science: a call for *Amaranthus* genomics. Pages 53–81 in Stewart Jr. CN ed Weedy and Invasive Plant Genomics. Ames, IA: Blackwell
- VanGessel M (2001) Glyphosate-resistant horseweed from Delaware. *Weed Sci* 49: 703-705
- Vila-Aiub MM, Balbi MC, Distéfano AJ, Fernández L, Hopp E, Yu Q, Powles SB (2012) Glyphosate resistance in perennial *Sorghum halepense* (Johnsongrass), endowed by reduced glyphosate translocation and leaf uptake. *Pest Manag Sci* 68: 430-436
- Vila-Aiub MM, Goh SS, Gaines TA, Han H, Busi R, Yu Q, Powles SB (2014) No fitness cost of glyphosate resistance endowed by massive EPSPS gene amplification in *Amaranthus palmeri*. *Planta* 239: 793-801
- Wakelin AM, Preston C (2006) A target-site mutation is present in a glyphosate-resistant *Lolium rigidum* population. *Weed Res* 46: 432-440
- Wiersma S, Chisholm S, Godar A, Stahlman P, Leach J, Westra P (2012) Gene amplification of EPSP synthase in glyphosate resistant *Kochia scoparia*. *Western Soc of Weed Sci Proc* 65:128
- Yu Q, Cairns A, Powles S (2007) Glyphosate, paraquat and ACCase multiple herbicide resistance evolved in a *Lolium rigidum* biotype. *Planta* 225: 499-513

CHAPTER 2

Association of EPSPS Gene Amplification with Glyphosate Resistance in Illinois

Waterhemp (*Amaranthus tuberculatus*) Populations

2.1 Abstract

With the frequency of glyphosate-resistant waterhemp increasing throughout the Midwest, the identification of resistant populations has become important for managing this species. However, high-throughput screening for glyphosate resistance in the greenhouse is tedious and inefficient. Research was conducted to document the occurrence of glyphosate-resistant waterhemp throughout the state of Illinois, and to determine whether a molecular assay for 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS) gene amplification can be used as an alternative means to detect resistant populations. Populations throughout the state of Illinois were collected and screened for glyphosate resistance using a whole-plant assay in a greenhouse, and survivors were examined for EPSPS gene amplification. Of 80 populations screened, 22 were glyphosate-resistant based on the greenhouse screen, and gene amplification was identified in 20 (91%) of the resistant populations. Although there are multiple mechanisms for glyphosate resistance in waterhemp, a molecular test for EPSPS gene amplification provides a rapid alternative for identification of glyphosate resistance in most populations.

2.2 Introduction

Glyphosate-resistant (GR) crops were first commercialized in 1996 and, within a decade, 96% of all soybean [*Glycine max* (L.) Merr.] grown in the United States were GR, as well as 61% of all maize (*Zea mays* L.) (Dill et al. 2008). The intensive use of glyphosate due to the

widespread adoption of GR crops increased selection for GR weeds. In 2000, horseweed [*Conyza canadensis* (L.) Cronq.] in GR soybean was the first reported weed to evolve resistance to glyphosate in a GR crop (VanGessel 2001). To date, 28 weed species worldwide have evolved resistance to glyphosate (Heap 2014).

The first documented GR waterhemp population was identified in a Missouri field with a history of continuous GR soybean and multiple glyphosate applications per year (Legleiter and Bradley 2008). GR waterhemp has now spread throughout Missouri (Rosenbaum and Bradley 2013; Schultz et al. 2014), Illinois (Riggins et al. 2012), and much of the central United States (Heap 2014). An investigation of the spread of glyphosate resistance from a single resistant population in Illinois suggests that independent selection events and long distance dispersal may be responsible for GR waterhemp evolution and spread (Liu et al. 2010).

Glyphosate targets 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS), an enzyme present in the chloroplasts of plants and in some bacteria. By blocking EPSPS activity, glyphosate ultimately prevents the synthesis of aromatic amino acids and all downstream products of the shikimic acid pathway. Currently identified mechanisms of weed resistance to glyphosate include reduced translocation (Lorraine-Colwill et al. 2003) most likely due to vacuolar sequestration (Ge et al. 2010), EPSPS point mutations (Baerson et al. 2002; Wakelin and Preston 2006) and EPSPS gene amplification (Gaines et al. 2010; Salas et al. 2012).

Amplification of the EPSPS gene was first identified as a GR mechanism in Palmer amaranth [*Amaranthus palmeri* S. Wats.], a weed species closely related to waterhemp (Gaines et al. 2010). While the mechanism of EPSPS gene amplification remains unknown, it is believed to be DNA-mediated, as intron sequences are maintained in the amplified copies (Gaines et al. 2013). Preliminary reports indicated that EPSPS gene amplification is present in waterhemp, but

the extent to which it contributes to glyphosate resistance was unclear (Bell et al. 2009; Shaner et al. 2011; Tranel et al. 2011). Subsequent reports have confirmed that gene amplification is associated with glyphosate resistance in waterhemp but also indicated that other mechanisms exist (Chatham et al. 2013). In addition to gene amplification, a Pro106Ser point mutation and reduced glyphosate translocation were found to confer resistance in some waterhemp populations (Bell et al. 2013; Nandula et al. 2013).

Identification of resistant populations can help growers make more informed management decisions. However, high-throughput screening for resistance in the greenhouse is an inefficient use of time and resources. The primary objective of this research was to determine if EPSPS gene amplification could be used as a proxy for glyphosate resistance in waterhemp. This study also was conducted to determine the distribution of GR waterhemp in Illinois. Waterhemp populations were obtained from counties throughout Illinois and screened for glyphosate resistance in the greenhouse. Surviving plants were evaluated to determine what percentage of resistant populations had EPSPS gene amplification.

2.3 Materials and Methods

2.3.1 Seed Collection

Seeds were collected from one or more agronomic fields in each of 80 Illinois counties during late summer in 2010 to obtain a wide representation of waterhemp populations in the state. Our objective was to obtain GR populations; therefore, rather than randomly sampling fields, we targeted fields in which GR waterhemp was suspected. Consequently, soybean fields were preferentially sampled over cornfields (since glyphosate is more likely to be the primary or sole herbicide used in the former), and waterhemp plants that likely received an herbicide

application were sampled (e.g., plants that were growing in a sprayer skip or that may have emerged after application were avoided).

From each field sampled, inflorescences from multiple female plants (typically five) were collected and pooled. Inflorescences were allowed to dry and seeds were threshed and stored at 4 C. Seeds obtained from a single field were designated as a population, and one population per county was arbitrarily selected for analysis. Prior to resistance screening, a subsample of seeds from each of the selected 80 populations was stratified to break seed dormancy. For stratification, seeds were first sterilized for 10 minutes in a 1:1 commercial bleach : deionized water solution. The seeds were then washed twice with deionized water, suspended in a 0.1% agarose solution, and stored at 4 C for at least 6 weeks prior to germination.

2.3.2 Resistance Testing

Germination, Planting, and Growth of Waterhemp

Seeds were germinated in standard sterile petri dishes lined with filter paper and moistened with deionized water. Petri dishes were sealed with Parafilm (Pechiney Plastic Packaging, Menasha, WI 54952) to prevent evaporation and placed in a germination chamber (Conviron, 590 Berry St., Winnipeg, Manitoba, Canada R3H 0R9) set to 12-hr days at 35 C and 12-hr nights at 15 C. After 48 hours, seedlings of uniform size were transplanted into containers (3.8 cm top diameter by 21 cm deep; Stuewe and Sons Inc., 31933 Rolland Dr, Tangen, OR 97389) filled with a 3:1:1:1 mixture of commercial potting mix (Sunshine Mix #1 / LC1, Sun Gro Horticulture, 770 Silver Street, Agawam, MA 01001) : soil : peat : sand. Plants were grown in a greenhouse set to 16-hr days with supplemental lighting to maintain a minimum of 800 $\mu\text{mol m}^{-2} \text{sec}^{-1}$. Greenhouse experiments were not performed during the winter months when day

length was short; even with supplemental lighting plants often flowered early and glyphosate activity was inconsistent. Temperature in the greenhouse was kept at 28 to 30 C during the day and 24 to 26 C at night. Plants were fertilized as needed with a solid slow-release fertilizer (Scott's Osmocote Classic 13-13-13, The Scotts Company, 1411 Scottslawn Rd., Marysville, OH 43041).

Herbicide Application and Resistance Evaluation

Fourteen plants of uniform size (4 to 5 cm tall) were chosen from each population and treated with 1260 g ae ha⁻¹ glyphosate (Roundup WeatherMax, Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, MO 63167). A pool of waterhemp seed collected from multiple Illinois counties in 2003 was used as a glyphosate-sensitive control, and the MO1 population (Legleiter and Bradley 2008) was used as a glyphosate-resistant control. Plants from both negative and positive controls were included in each spray pass. Herbicide application was carried out using a moving-nozzle cabinet spray chamber equipped with a 80015 even flat fan nozzle (TeeJet Technologies, P.O. Box 7900, Wheaton, IL 60187) held approximately 46 cm above the plant canopy. The spray chamber was calibrated to deliver 187 L ha⁻¹.

At 14 DAT, individual plants were visually evaluated and rated as sensitive (no green tissue), intermediate (some green tissue but little or no new growth) or resistant (abundant new growth). A population was considered resistant if two or more resistant individuals were present among the 14 individuals, or if one resistant individual and two or more intermediate individuals were present. The experiment was repeated on all populations identified as resistant. Tissue samples were taken from plants categorized as resistant for molecular investigation of resistance

mechanisms. Tissue samples consisted of a single, newly emerging leaf, approximately 1 to 2 cm long.

2.3.3 Examination of Resistance Mechanisms

Sample Preparation

DNA was extracted using the hexadecyltrimethylammonium bromide (CTAB) method previously described by Doyle and Doyle (1990) and was examined for quality and quantity using a spectrophotometer (NanoDrop 1000 Spectrophotometer, Thermo Fisher Scientific, 81 Wyman St., Waltham, MA 02451). Samples were diluted to 10 ng μL^{-1} for all downstream applications.

EPSPS Gene Amplification

Quantitative real-time polymerase chain reaction (qPCR) was used to determine relative EPSPS gene copy number compared to a one-copy reference gene using methods described previously (Délye et al. 2014; Ma et al. 2013).

Alternative Target-Site Resistance

Resistant plants were screened for the point mutations resulting in the Pro106Ser substitution (Bell et al. 2013; Nandula et al. 2013) using a derived cleaved amplified polymorphic sequences (dCAPS) assay designed according to procedures described by Délye et al. (2014). A fragment containing codon 106 was amplified using the forward primer EPSf1 (5'-ATG TTG GAC GCT CTC AGA ACT CTT GGT-3') originally designed for qPCR (Gaines et al. 2010) and reverse primer eps106wt-R3 (5'-CTC CAG CAA CGG CAA CCG CAA CTG

TCC ATG-3'). The reverse primer was engineered to create a *NcoI* restriction site by substituting a cytosine for an adenine at the fourth nucleotide from the 3' end of the primer. The resulting PCR product was subjected to a restriction enzyme digest with *NcoI* (New England BioLabs Inc., 240 County Road, Ipswich, MA 01938) and then fractionated on a 2% agarose gel containing 0.5 $\mu\text{g mL}^{-1}$ ethidium bromide. Gels were visualized with ultraviolet light and results were compared to samples from known glyphosate-susceptible waterhemp plants.

EPSPS Gene Sequencing

A subset of samples with the Pro106Ser point mutation identified using the dCAPS assay was sequenced to confirm the accuracy of the assay. PCR products from primers EPSF1 and eps106wt-R3 were visualized with gel electrophoreses as described above to confirm the presence of the correct amplicon. The remaining PCR product was purified (E.Z.N.A. Cycle Pure Kit, Omega Bio-Tek, Inc., 400 Pinnacle Way, Suite 450, Norcross, GA 30071) and used in a sequencing reaction (BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems Inc., 850 Lincoln Centre Drive Foster City, CA 94404) using the EPSF1 primer. The resulting product was analyzed by the W.M. Keck Center for Comparative and Functional Genomics using an AB 3730xl DNA analyzer (Applied Biosystems Inc.). EPSPS gene sequences were aligned to waterhemp sequences from glyphosate-susceptible lines in GenBank (FJ869881 and FJ869880) using MEGA6 (Tamura et al. 2013).

2.4 Results and Discussion

From 80 populations sampled throughout Illinois in 2010, 22 (28%) were identified as glyphosate resistant. All 22 populations were verified as resistant during the second screen. Our

observed frequency should not be taken as the probability that any one field in Illinois contains GR waterhemp because: 1) sampling was non-random and biased towards suspect GR populations; 2) the frequency of resistance varied across the state, with the majority of resistant populations coming from the southern part of the state (Figure 2.1); and 3) sampling was performed in 2010 and the frequency of GR waterhemp has likely increased.

Because fields were targeted in which GR waterhemp was suspected, we were surprised that such a low percentage of the populations were, in fact, glyphosate resistant. That less than a third of the populations were confirmed as glyphosate resistant suggests that the appearance of waterhemp in fields near the end of the growing season more often than not was due to management practices rather than to glyphosate resistance. Alternatively, the low frequency of resistant populations found in this study may be due in part to the relatively conservative set of criteria used to differentiate between sensitive and resistant populations. These strict criteria were used to reduce the risk of falsely identifying resistant populations, but it may have excluded populations with low frequencies or low magnitudes of resistance. Our results are in contrast to those reported by Riggins et al. (2012), who found that at least two-thirds of waterhemp populations suspected of glyphosate resistance were truly resistant. In their study, however, growers – who had site history and management information on which to base their suspicions – submitted the samples. As we found in the present study, identifying GR waterhemp populations without such information is difficult. Rosenbaum and Bradley (2013) showed that signs of herbicide survival, such as stunting or excessive branching resulting from loss of the apical meristem, were strong indicators of glyphosate resistance. We did not use this criterion during our sampling.

The majority of waterhemp populations, 86%, were collected in soybean fields, while only 14% were collected in cornfields (Figure 2.1). All populations identified as resistant were collected from soybean fields. While this may be simply an artifact of the unequal collection of populations from corn and soybean fields, it may indicate that waterhemp escapes in soybean fields are more likely to be glyphosate resistant. There are fewer postemergence herbicide options to control waterhemp in soybean than in corn and, therefore, reliance on glyphosate in soybean is likely to be higher than in corn.

Survivors from each of the resistant populations were screened for EPSPS gene amplification using qPCR. A threshold relative copy number was chosen based on a distribution of relative EPSPS gene copy numbers from sensitive controls used on each plate of qPCR in this study. The average copy number for sensitive controls was 0.98 and ranged from 0.60 to 1.87. The copy numbers from the sensitive controls were reciprocally transformed to normalize the data, and the copy number that was greater than 95% of the sensitive control values was rounded to the nearest tenth and set as the threshold. Samples with a copy number higher than this threshold, 1.4, were considered to have elevated EPSPS gene copy number. Results indicated that 91% of resistant populations had at least one survivor with gene amplification. Of these populations, all but two (from Saline and Pulaski Counties) were verified as having EPSPS gene amplification in the second screen. These two populations both had a relatively small proportion of plants with elevated EPSPS gene copy number in the first screen (Figure 2.2). Alternative resistance mechanisms may explain the discrepancy between the first and second screen. Only two resistant populations, from Schuyler and Mason counties, did not have elevated EPSPS gene copy number based on the threshold value used. This comprised only 9% of the resistant populations studied (Figure 2.2). These results are consistent with previous findings that gene

amplification is associated with glyphosate resistance in multiple geographic locations (Chatham et al. 2012; Chatham et al. 2013).

Of the resistant populations with EPSPS gene amplification, 7 had elevated copy number in every resistant plant sampled (Figure 2.2). The remaining 13 populations had elevated copy numbers in varying proportions of the total plants sampled. While it is possible that this is due to escapes (i.e., some of the survivors were not resistant), it is likely that some of the populations have an additional mechanism(s) of resistance. Chatham et al. (2012) also found a number of GR waterhemp plants without gene amplification.

The average copy number of the plants with copy number above the threshold in each population was determined to look at the variation in copy number magnitude among populations. With the exception of the populations from Morgan and Bond counties, all populations had an average copy number above 3.0, with the highest average copy number at 8.6 in the Lawrence County population. Morgan and Bond populations each had a relatively low proportion of plants with elevated copy number as well as a low average copy number among those that breached the threshold value. It is possible that these populations may not truly have EPSPS gene amplification; the plants with elevated copy number in these populations may fall within the 5% of sensitive plants with copy number above the set threshold value. Given the uncertainty of the presence of gene amplification in these two populations, a more conservative estimate of the false negative rate when using gene amplification as a proxy for glyphosate resistance would be 18%.

Further investigation of alternative mechanisms of resistance was carried out using a dCAPS assay designed to detect the Pro106Ser substitution resulting from a cytosine to thymine nucleotide change previously identified in waterhemp (Bell et al. 2013; Nandula et al. 2013). The

assay employs the use of a restriction enzyme, *NcoI*, which cuts an amplified region of EPSPS without the mutation (i.e., the wild-type), producing a smaller DNA fragment that migrates faster during agarose gel electrophoresis. Since this assay recognizes the wild type genotype, it is possible that alternative amino acid substitutions resulting from a change to the first nucleotide of the proline codon, including alanine and threonine may also be recognized. Samples from the two populations without gene amplification, as well as samples from populations in which gene amplification was not present in all samples were tested using the assay. The Pro106Ser point mutation was found in the majority of samples from the Schuyler County population (Figure 2.3). All samples with the mutation were heterozygous, but this is not surprising given that only 43% of the Schuyler population survived glyphosate treatment. Using Hardy-Weinberg proportions (Andrews 2010), it can be determined that the EPSPS allele with serine at amino acid 106 will be mostly present in heterozygotes, with only about 1 in 17 plants expected to be homozygous for this allele. The Pro106Ser mutation was also found at a low frequency in several other populations. In the population from Lawrence County, the mutation was found in several resistant plants without gene amplification. The Pro106Ser mutation also showed up at a low frequency in GR individuals from the Madison, Jackson, and Morgan County populations. A subset of samples from the populations with the Pro106Ser mutation was chosen for sequencing, which confirmed the accuracy of the assay (Figure 2.3). Although the Pro106Ser mutation does appear to be associated with resistance, questions remain regarding the ability of the mutation to confer resistance. While Nandula et al. (2013) consistently found the mutation in all GR samples, Bell et al. (2013) observed mortality of some plants with the Pro106Ser mutation at higher glyphosate rates. Further investigation of the interplay between multiple mechanisms of resistance in the populations studied here may be worthwhile. For example, the discovery of

amplified EPSPS containing the Pro106Ser point mutation could provide information regarding the evolution of these resistance mechanisms.

The results herein suggest that a third mechanism of glyphosate resistance may be acting in Illinois waterhemp populations. Assuming EPSPS gene amplification is the primary mechanism of resistance in Illinois waterhemp populations, the majority of resistant populations should have elevated EPSPS copy number in all resistant individuals. However, elevated EPSPS copy number was not found in all resistant individuals for the majority of the resistant populations with gene amplification (Figure 2.2). This suggests that the majority of the resistant populations found in Illinois may have multiple mechanisms of glyphosate resistance. Nandula et al. (2013) found non-target site resistance in a waterhemp population from Mississippi. Altered translocation and absorption is the only reported mechanism of glyphosate resistance not studied herein. Non-target-site-based mechanisms may be able to explain the resistant populations that did not have elevated copy number or the Pro106Ser mutation, as well as the large number of populations that had gene amplification present in some, but not all, resistant individuals. Translocation studies in addition to dose responses and genetic inheritance studies would be necessary to identify any alternative glyphosate resistance mechanisms working in these populations.

This study suggests that EPSPS gene amplification is the primary mechanism of glyphosate resistance in waterhemp populations from Illinois, but that other mechanisms exist. While a more thorough, multi-year survey of glyphosate resistance throughout Illinois is necessary to more accurately describe the distribution and spread of GR waterhemp, the current study provides data to support the use and accuracy of various molecular tools for the identification of GR waterhemp in such studies.

2.5 Acknowledgements

We thank Kevin Bradley from the University of Missouri for providing the MO1 population. The USDA National Institute of Food and Agriculture (Hatch project ILLU-802-923) and the Illinois Soybean Association and Soybean Checkoff provided partial funding of this research.

2.6 Literature Cited

- Andrews C (2010) The Hardy-Weinberg principle. *Nature Education Knowledge* 3:65
- Baerson SR, Rodriguez DJ, Tran M, Feng Y, Viest NA, Dill GM (2002) Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiol* 129:1265-1275
- Bell MS, Tranel PJ, Riggins CW (2009) Glyphosate resistance in waterhemp: inheritance and EPSPS copy number. *North Central Weed Sci Soc Proc* 64:100
- Bell MA, Hager AG, Tranel PJ (2013) Multiple resistance to herbicides from four site-of-action groups in waterhemp (*Amaranthus tuberculatus*). *Weed Sci* 61:460-468
- Chatham LA, Riggins C, Owen MD, Tranel P (2012) Association of EPSPS gene amplification with glyphosate resistance in waterhemp. *North Central Weed Sci Soc Proc* 67:50
- Chatham LA, Riggins CW, Martin JR, Kruger GR, Bradley KW, Peterson DE, Jugulam M, Tranel P (2013) A multi-state study of the association between glyphosate resistance and EPSPS gene amplification in waterhemp. *North Central Weed Sci Soc Proc* 68:127
- Délye C, Duhoux A, Pernin F, Riggins CW, Tranel PJ (2014) Research methods in weed science: molecular mechanisms of herbicide resistance. *Weed Sci* 62: In press

- Dill GM, Jacob CA, Padgett SR (2008) Glyphosate-resistant crops: adoption, use and future considerations. *Pest Manag Sci* 64: 326-331
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc Natl Acad Sci U.S.A.* 107:1029-1034
- Gaines TA, Wright AA, Molin WT, Lorentz L, Riggins CW, Tranel PJ, Beffa R, Westra P, Powles SB (2013) Identification of genetic elements associated with EPSPS gene amplification. *PloS one* 8(6): e65819
- Ge X, d'Avignon DA, Ackerman JJH, Sammons RD (2010) Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. *Pest Manag Sci* 66:345–348
- Heap IM (2014) International Survey of Herbicide Resistant Weeds. www.weedscience.org. Accessed May 13, 2014
- Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Sci* 56:582-587
- Liu J, Tranel PJ, Davis AS (2010) Modeling the spread of glyphosate resistant waterhemp. *Proc North Central Weed Sci. Soc.* 65:59
- Lorraine-Colwill DF, Powles SB, Hawkes TR, Hollinshead PH, Warner SAJ, Preston C, (2003) Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. *Pestic Biochem Physiol* 74:62-72

- Ma R, Kaundun SS, Tranel PJ, Riggins CW, McGinness DL, Hager AG, Hawkes T, McIndoes E, Riechers DE (2013) Distinct detoxification mechanisms confer resistance to mesotrione and atrazine in a population of waterhemp. *Plant Physiol* 163:363-377
- Nandula VK, Ray JD, Ribeiro DN, Pand Z, Reddy KN (2013) Glyphosate resistance in tall waterhemp (*Amaranthus tuberculatus*) from Mississippi is due to both altered target-site and nontarget-site mechanisms. *Weed Sci* 61:374-383
- Riggins C, Hager AG, Tranel P (2012) Three years of testing Illinois waterhemp populations for multiple resistance to glyphosate, PPO inhibitors, and ALS inhibitors. *North Central Weed Sci Soc Proc* 67:122
- Rosenbaum KK, Bradley KW (2013) A survey of glyphosate-resistant waterhemp (*Amaranthus rudis*) in Missouri soybean fields and prediction of glyphosate resistance in future waterhemp populations based on in-field observations and management practices. *Weed Technol* 27:656-663
- Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW, Scott RC, Burgos NR (2012) EPSPS gene amplification in glyphosate-resistant Italian Ryegrass (*Lolium perenne* ssp. *Multiflorum*) from Arkansas, USA. *Pest Manag Sci* 68:1223-1230
- Schultz J, Riley EB, Wait JD, Bradley KW (2014) Distribution of multiple herbicide resistance in Missouri waterhemp populations. *Weed Sci Soc Am Proc* 54:290
- Shaner DL, Lindenmeyer RB, Ostlie MH (2011) What have the mechanism of resistance to glyphosate taught us? *Pest Manag Sci* 68:3-9
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30:2725-2729

Tranel PJ, Riggins CW, Bell MS, Hager AG (2011) Herbicide resistances in *Amaranthus tuberculatus*: a call for new options. J Agric Food Chem 59:5808-5812

VanGessel M (2001) Glyphosate-resistant horseweed from Delaware. Weed Sci 49: 703-705

Wakelin AM, Preston C (2006) A target-site mutation is present in a glyphosate-resistant *Lolium rigidum* population. Weed Res 46:432-440

2.7 Figures

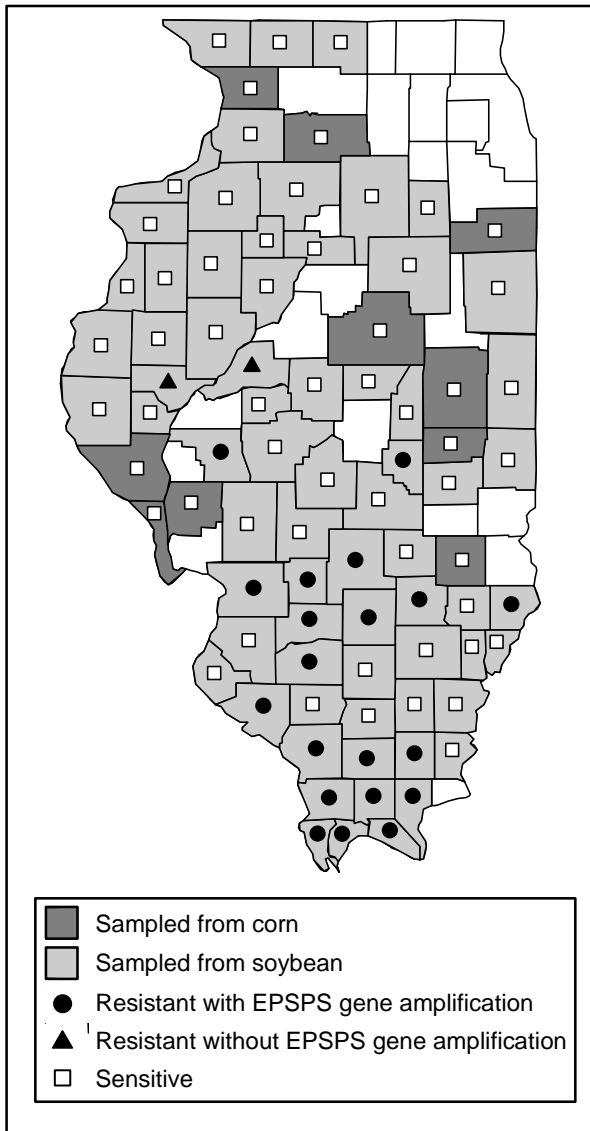


Figure 2.1: Map of Illinois counties from which populations were collected for screening.

County colors refer to the crop planted in the field where the population was collected; symbols within each county indicate the resistance screening result for each population.

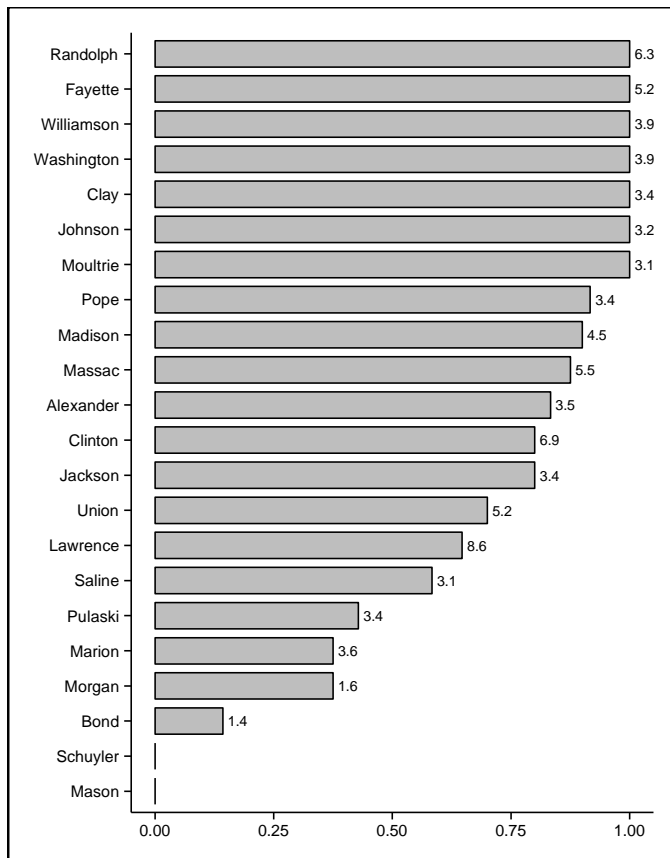


Figure 2.2: Proportion of plants with EPSPS gene amplification from glyphosate-resistant populations. Numbers to the right of each bar refer to the mean copy number of the plants with gene amplification in the population (i.e., plants within the population that did not have gene amplification were excluded from the mean calculation). Populations from Schuyler County and Mason County had no individuals with EPSPS gene amplification.

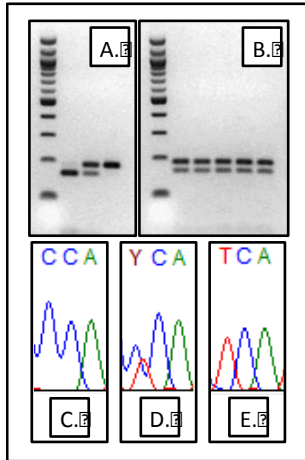


Figure 2.3: (A) Results of the dCAPS assay for the identification of the Pro106Ser point mutation in waterhemp showing samples of each possible genotype: homozygous with two proline alleles, heterozygous with one proline and one serine, and homozygous with two serine alleles. (B) Assay results from a subset of Schuyler County population samples, each of which was heterozygous with one proline allele and one serine allele. (C-E) Chromatograms showing the nucleotide sequence of codon 106 in each of the genotypes illustrated in (A): proline/proline (C), proline/serine (D), serine/serine (E).

CHAPTER 3

A Multi-State Study of the Association Between Glyphosate Resistance and EPSPS Gene Amplification in Waterhemp (*Amaranthus tuberculatus*)

3.1 Abstract

Waterhemp has become an increasingly problematic weed in the Midwest, having now evolved resistances to herbicides from six different site-of-action groups. Glyphosate-resistant waterhemp in the Midwest is especially concerning given the economic importance of glyphosate in corn and soybean production. Amplification of the target-site gene, EPSPS (5-enolpyruvylshikimate-3-phosphate synthase) was found to be the mechanism of resistance in Palmer amaranth, a species closely related to waterhemp. Here, the relationship between glyphosate resistance and EPSPS gene amplification in waterhemp was investigated. Glyphosate dose responses were performed at field sites with glyphosate-resistant waterhemp in Illinois, Kansas, Kentucky, Missouri and Nebraska, and relative EPSPS copy number of survivors was determined via qPCR. Waterhemp control increased with increasing glyphosate rate at all locations but no population was completely controlled even at the highest rate (3360 g ae ha⁻¹). EPSPS gene amplification was present in plants from four of five locations (IL, KS, MO and NE) and the proportion of plants with elevated copy number was generally higher in survivors from glyphosate-treated plots than in plants from the untreated control plots. Copy number magnitude varied by site, but an overall trend of increasing copy number with increasing rate was observed in populations with gene amplification, suggesting that waterhemp plants with more EPSPS copies are more resistant. Survivors from the KY population did not have elevated EPSPS copy number. Instead, resistance in this population was attributed to the EPSPS

Pro106Ser mutation. Results herein show a quantitative relationship between glyphosate resistance and EPSPS gene amplification in some waterhemp populations while highlighting that other mechanisms confer glyphosate resistance in waterhemp as well.

3.2 Introduction

Glyphosate was first commercialized in 1974 and has since been considered one of the most ideal herbicides. It has broad-spectrum control, good translocation combined with a slow mode of action, and low mammalian and environmental toxicity (Duke and Powles 2008). The addition of crop selectivity to this list endowed by glyphosate-resistant (GR) crop technology commercialized in 1996 has made glyphosate arguably the most economically important herbicide worldwide. However, the pandemic use of glyphosate that followed has led to intense selection pressure on weeds to evolve resistance. To date, 28 weeds across 24 countries have evolved resistance to glyphosate (Heap 2014).

Few species have invaded agronomic fields and developed such a widespread distribution as quickly or efficiently as waterhemp. This success is due in part to waterhemp's dioecious strategy of sexual reproduction, C₄ photosynthetic pathway, prolonged emergence period, and propensity to evolve resistance (Tranel and Trucco 2009). In a production field setting, waterhemp competition can reduce soybean yields up to 43% (Hager et al. 2002), and season long competition in corn can reduce yields as much as 74% (Steckel and Sprague 2004). Glyphosate was the fourth of six herbicides with distinct sites of action to which waterhemp has evolved resistance. The first glyphosate-resistant waterhemp population was identified in a Missouri field with a history of continuous GR soybean and multiple yearly glyphosate applications (Legleiter and Bradley 2008).

Glyphosate inhibits EPSPS (enol-5-pyruvyl-shikimate-3-phosphate synthase), a key enzyme in the shikimic acid pathway. This inhibition of the shikimic acid pathway prevents the synthesis of aromatic amino acids and all downstream products. Blocking normal EPSPS activity has also been shown to reduce feedback inhibition in the shikimic acid pathway. This deregulation could result in a massive carbon flux away from other important cellular pathways to the shikimic acid pathway (Duke and Powles 2008). Weeds have developed a number of resistance mechanisms in an attempt to thwart glyphosate. Currently identified mechanisms of glyphosate resistance include reduced translocation (Lorraine-Colwill et al. 2003) which in most cases is likely due to vacuolar sequestration (Ge et al. 2010), target-site point mutations at Pro106 (Baerson et al. 2002; Wakelin and Preston 2006) and EPSPS gene amplification (Gaines et al. 2010; Salas et al. 2012; Wiersma et al. 2012).

EPSPS gene amplification was first discovered in a weed closely related to waterhemp, Palmer amaranth (Gaines et al. 2010). Multiple studies in Palmer amaranth have demonstrated the positive correlation between glyphosate resistance and EPSPS genomic copy number, EPSPS transcript and protein expression, and EPSPS enzyme activity (Gaines et al. 2010; Ribeiro et al. 2014). Additional EPSPS proteins resulting from multiple EPSPS gene copies allow the plant to survive in the presence of glyphosate since the herbicide becomes overwhelmed by a massive amount of target-site protein (Powles 2010). While the mechanism of gene amplification remains unknown, MITEs (miniature inverted-repeat transposable elements), a putative *Ac (Activator)* transposase, and repetitive sequence region were found to be associated with amplified EPSPS copies in GR Palmer amaranth individuals. Introns were also confirmed to be present in amplified gene copies, providing further evidence for a DNA-mediated mechanism of gene amplification (Gaines et al. 2013). Preliminary reports suggest that EPSPS gene amplification

may be associated with glyphosate resistance in waterhemp as well, (Bell et al. 2009; Tranel et al. 2011; Shaner et al. 2011). However, the extent to which elevated EPSPS copy number contributes to glyphosate resistance remains unclear. In addition to EPSPS gene amplification as a potential mechanism of resistance, a Pro106Ser point mutation (Bell et al. 2013; Nandula et al. 2013) and reduced glyphosate translocation were also found to confer resistance in some waterhemp populations (Nandula et al. 2013).

While preliminary reports have linked glyphosate resistance and EPSPS gene amplification in waterhemp, the relationship between EPSPS copy number magnitude and glyphosate resistance under field conditions has not been examined. Is EPSPS gene amplification the primary GR mechanism in geographically diverse waterhemp populations? Do higher glyphosate rates select for plants with higher EPSPS gene copy numbers? With these questions in mind, the primary objective of this study was to determine EPSPS gene copy numbers in waterhemp plants surviving different rates of glyphosate at multiple field locations in the Midwest.

3.3 Materials and Methods

3.3.1 Field Studies

Field studies were conducted in Douglas County, Illinois (IL), Franklin County, Kansas (KS), Hancock County, Kentucky (KY), Randolph County, Missouri (MO) and Dodge County, Nebraska (NE) at sites with suspected or confirmed GR waterhemp (Figure 3.1). Studies were set up as randomized complete block designs with 0x, 0.5x, 1x, 2x, and 4x rates of formulated glyphosate ($1x = 840 \text{ g ae ha}^{-1}$) in three or four replications. Rates were chosen based on the results of a pilot study performed on GR waterhemp populations in IL and Iowa (IA) the

previous year (Chatham et al. 2012). Herbicide applications were made in mid-June to mid-July 2013 when the majority of waterhemp plants were 8 to 12 cm in height. Studies were carried out using standard small plot research procedures using pressurized CO₂ backpack sprayers; however, specific application equipment and sprayer calibration varied by location. Plot sizes ranged from 10 to 20 m², and were planted with either soybean or no crop. Two to four weeks after treatment, counts and/or visual observations of glyphosate activity were taken and leaf samples were obtained from at least four survivors in each plot (replication) for each treatment at each location. Leaf samples were taken from newly emerging leaves approximately 1 to 2 cm in length, stored at 4 C and shipped on ice to the University of Illinois.

3.3.2 Examination of Resistance Mechanisms

Sample preparation

The CTAB (hexadecyltrimethylammonium bromide) DNA isolation method described previously by Doyle and Doyle (1990) was used to extract DNA from leaf samples. The concentration and purity of each sample was examined using a spectrophotometer (NanoDrop 1000 Spectrophotometer, Thermo Fisher Scientific, 81 Wyman St., Waltham, MA 02451) and each was diluted to 10 ng μL^{-1} for all subsequent procedures.

EPSPS Gene Amplification

Samples were tested for elevated EPSPS copy number compared to a one-copy reference gene using quantitative real-time PCR as described previously (Ma et al. 2013, Délye et al. 2014). A threshold EPSPS copy number value was set to the maximum relative copy number observed for any of the glyphosate-sensitive controls used in the study. Samples with a relative

EPSPS copy number above this threshold (1.5) were considered to have elevated copy number. All GR controls were consistently above the threshold.

Alternative Target-Site Resistance

Select samples without gene amplification were screened for the Pro106Ser mutation (Bell et al. 2013, Nandula et al. 2013) using a dCAPS assay designed according to methods described by Délye et al. (2014) and performed as described previously (Chatham et al. submitted).

EPSPS Sequencing

Several samples that tested positive for the Pro106Ser mutation using the dCAPS assay were sequenced to confirm the results. PCR was performed using primers EPSF1, originally designed for qPCR (Gaines et al. 2010), and eps106wt-R3, originally designed for use in the dCAPS assay mentioned above. After confirming the presence of the correct amplicon via agarose gel electrophoresis (1% agarose; 0.5 $\mu\text{g mL}^{-1}$ ethidium bromide), the PCR product was purified (E.Z.N.A. Cycle Pure Kit, Omega Bio-Tek, Inc., 400 Pinnacle Way, Suite 450, Norcross, GA 30071) and sequenced (BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems Inc., 850 Lincoln Centre Drive Foster City, CA 94404) with the EPSF1 primer. Products were further analyzed by the W. M. Keck Center for Comparative and Functional Genomics using an AB 3730xl DNA analyzer (Applied Biosystems Inc.). Returned sequences were aligned to waterhemp EPSPS sequences from glyphosate-susceptible lines in GenBank (FJ869881 and FJ869880) with MEGA6 (Tamura et al. 2013).

3.3.3 Statistical Analysis

Statistical analysis of the correlation between glyphosate rate and relative EPSPS copy number was carried out in R (v3.0.3) (R Core Team 2014) using bootstrapping of the correlation coefficient (r) with 99,999 resamples and treating location as a random effect. The resampled distributions were compared to the real copy number data to determine how many times the correlation coefficient from the original data set was observed by chance alone in the resampled distributions.

3. 4 Results and Discussion

3.4.1 Waterhemp Control

Dose response studies confirmed the presence of GR waterhemp at each field location. Complete control was not seen in any of the plots at any of the locations, even with the highest glyphosate rate. Because of the variability among locations in how the experiments were conducted, the dose-response data were not statistically analyzed. However, visual inspection of the data (Figure 3.2) reveals a general trend of increasing waterhemp control with increasing glyphosate rate at all locations. This pattern of control with glyphosate has been seen previously with GR waterhemp (Legleiter and Bradley 2008; Patton et al. 2012).

3.4.2 EPSPS Gene Amplification

Copy number analysis revealed that EPSPS gene amplification was present at four (IL, KS, MO, NE) of the five locations studied (Figure 3.3). When a pilot study was conducted in 2012, gene amplification was also observed at the same location in IL and at a location in IA (Chatham et al. 2012). Counting the IA location, EPSPS gene amplification was observed in five

of six locations investigated. Because the IA location was not included in 2013 (due to weather events) it is not discussed further, and the KY location, where EPSPS gene amplification was not observed, is discussed in a following section.

Several surviving plants without EPSPS gene amplification were found at each glyphosate rate at each location (Figure 3.3). If EPSPS gene amplification is solely responsible for resistance in these populations, one would expect to see elevated EPSPS copy number in all plants surviving glyphosate, particularly at the higher glyphosate rates. While some survivors may have been escapes and not truly resistant, their presence at all four locations – and even at the highest rate – is difficult to ignore. Alternative mechanisms of resistance are a more likely explanation for these anomalies.

To further examine whether EPSPS gene amplification is associated with resistance, a chi-square goodness of fit test was performed on the copy number data obtained from the four locations at which amplification was observed (Table 1). The proportion of plants with elevated copy number sampled from untreated control plots was used as the expected proportion and the proportion of plants with elevated copy number in all glyphosate treatments was used as the observed proportion. In IL, KS and NE, the proportion of surviving plants with elevated EPSPS copy number was significantly different from the untreated control ($p < 0.001$), confirming that gene amplification is associated with glyphosate resistance in waterhemp. However, in Missouri, the proportion of plants with elevated copy number in the glyphosate treated plots was not statistically different from what was expected based on the untreated control ($p = 0.386$).

Although not statistically significant, the MO population does show a general trend of increasing copy number with rate (Fig. 3). The lack of significance here may be due to the high background

frequency of plants with elevated copy number in the population (75% based on plants from control plots).

An examination of the raw EPSPS copy number data revealed that copy number magnitude may vary by location (Figure 3.2), and the combined data for all treatments at each location (Figure 3.4a) showed an obvious difference in the copy number distributions for each population. The populations from IL, KS and MO clearly had individuals with EPSPS gene amplification; the average copy number of plants with gene amplification at these locations was 4.1, 4.0, and 3.9, respectively. These averages are similar to those found in the original GR population (MO1) (Legleiter and Bradley 2008) in the first accounts of EPSPS gene amplification in waterhemp (Bell et al. 2009). However, the EPSPS copy number distribution from the NE population was distinct from those of the IL, KS and MO populations (Figure 3.4a). The magnitude of EPSPS copy number in NE was lower, with an average copy number of 2.1 among plants with elevated EPSPS copy number.

Differences in background frequencies of plants with high EPSPS copy numbers (e.g., the MO population relative to other populations) and in the copy number magnitudes (e.g., the NE population relative to the other populations) could be explained by temporal differences in the GR evolutionary process at each location. Anecdotal reports indicate that the waterhemp populations at the IL and KS locations were identified as GR in 2010, whereas the MO population was suspected to be resistant several years prior and the NE population was not known to be resistant until 2012. Because resistance evolution is a gradual process, and because of varying lag times between observation and reporting of resistance, it is impossible to accurately chronologically compare resistance evolution at the different locations. Nevertheless, our EPSPS copy number data are consistent with a longer and shorter GR evolutionary history in

the MO and NE populations, respectively. For example, longer-term selection for glyphosate resistance in the MO population may have given this population more time to accumulate the gene amplification mechanism as well as any alternative mechanism that might confound the correlation between resistance and copy number.

Among the four populations that did have EPSPS gene amplification, a general trend of increasing copy number with increasing glyphosate rate was observed in the raw copy number data (Fig. 2). This relationship is more clearly depicted when data are combined among populations with gene amplification (IL, KS, MO, and NE; Figure 3.4b). However an examination of the interaction between median relative EPSPS copy number and glyphosate rate for these four locations shows that, while this relationship exists, it seems to vary slightly by population (Figure 3.4c).

While a general trend of increasing copy number with increasing rate can be seen in the combined EPSPS copy number data (Figure 3.4b) and individually by location (Figures 3.3 and 3.4c.), statistical characterization of the trend was not straightforward: the data set did not meet the assumptions required to perform parametric statistics. Both the original data and the linear model residuals were not normally distributed (Shapiro-Wilk $p=2.2e-16$, $p=3.2e-14$, respectively) and the variances were not equal among locations (Bartlett's test, $p=2.2e-16$) or treatments (Bartlett's test $p=3.9e-6$). These violations of the parametric ANOVA assumptions were also obvious from visual inspection of the data (Figures 3.4a and 3.4b). Combined with the unequal sample sizes, these results warranted a nonparametric statistical approach to analyze the data. A significant correlation was found between glyphosate rate and EPSPS copy number ($r=0.3$, $p<0.001$, 99999 resamples) in the combined data from the populations with EPSPS gene amplification. These results indicate that gene amplification is associated with glyphosate

resistance in waterhemp and suggests that higher copy number survivors were present at higher rates. To further explore this hypothesis, bootstrapping was performed again on the combined data, without the data from control treatments in each state. The increase in copy number between the control and the first glyphosate treatment could have contributed largely to the observed r value. Without the 0x data, a smaller yet still significant correlation was observed ($r=0.20$, $p<0.001$, 99999 resamples), confirming that a quantitative relationship between EPSPS copy number and glyphosate rate exists in these populations. These results indicate that higher glyphosate rates select for individuals with higher average EPSPS copy number, suggesting that plants with more EPSPS copies have a higher level of resistance than those with fewer copies. Gaines et al. (2010) suggested that a similar pattern of increasing EPSPS copy number with increasing glyphosate rate exists in Palmer amaranth.

3.4.3 KY Population

There is a clear connection between EPSPS gene amplification and glyphosate resistance in several of the populations studied; however, there is also evidence for alternative resistance mechanisms. Since none of the survivors from the KY population had elevated copy number, the samples were screened with a dCAPS assay designed to detect the Pro106Ser mutation, which has previously been found to be associated with glyphosate resistance in waterhemp (Bell et al. 2013, Nandula et al. 2013). This assay detects the wild type genotype and it is therefore possible that alternative substituting amino acids (alanine or leucine) also resulting from a change to the first nucleotide of the proline codon would be recognized. The Pro106Ser mutation was present in the KY population, and found in 69% percent of samples from the control plots. However, only 6% of plants from the control were homozygous for the mutation, with both alleles having

the serine substitution. At the 2x and 4x rates, 100% of survivors had the mutation, and at the 4x rate, 67% were homozygous for the mutation (Figure 3.5). The excellent relationship between glyphosate rate and the proportion of survivors with the mutation implies that the Pro106Ser point mutation is responsible for glyphosate resistance in the KY population. Individuals that survived glyphosate treatment but did not have elevated EPSPS copy number in the other locations were also examined, but the Pro106Ser mutation was not found in any of the samples. Several samples from the KY population of each genotype were sequenced and results confirmed that the Pro106Ser mutation was present and that the dCAPS assay was working correctly (data not shown).

3.4.4 Alternative Mechanisms

The only other known potential mechanism of glyphosate resistance in waterhemp not examined herein is altered translocation and/or uptake (Nandula et al. 2013). Examination of this mechanism would require whole plant specimens of the glyphosate survivors to examine. It is possible that altered translocation is present in some of the populations studied here. This might explain why some plants without gene amplification survived glyphosate treatment, especially at the higher rates. Multiple mechanisms of resistance in some of these populations may have confounded the results making examination of the relationship between EPSPS gene amplification and glyphosate resistance more difficult. Therefore, the correlation between rate and copy number may be underestimated here compared to a more homogeneous population with gene amplification as the only existing mechanism of glyphosate resistance. The potential for multiple resistance mechanisms may also have contributed to the lack of a significant chi-square goodness of fit p-value in the MO population.

In conclusion, results from this study confirm that an association between glyphosate resistance and EPSPS gene amplification exists in some waterhemp populations. Gene amplification was found in five of six populations studied (including the IA population in the 2012 pilot study). Given the locations of the populations studied, EPSPS gene amplification appears to be a common and widespread mechanism of resistance (Figure 3.1). Survey studies carried out in Illinois and Missouri both found gene amplification present in the majority of GR waterhemp populations (Chatham et al. submitted, Schultz et al. submitted). Among populations with elevated EPSPS copy number, a positive correlation between copy number and level of resistance was observed. Copy number magnitude varied by location, but overall was significantly lower than copy numbers observed in Palmer amaranth. Not all the populations studied had gene amplification; the Pro106Ser mutation was found to be primarily responsible for resistance in the KY population. While gene amplification appears to be the primary mechanism of resistance in waterhemp, it is clear that other mechanisms exist. Further investigation of these mechanisms and their interplay when combined may be required to completely understand glyphosate resistance in waterhemp.

3.5 Acknowledgements

We thank Aaron Hager and Doug Maxwell for assistance in carrying out the Illinois dose response study, and Chance Riggins for assistance with qPCR. The USDA National Institute of Food and Agriculture (Hatch project ILLU-802-923) provided partial funding for this research.

3.6 Literature Cited

- Baerson SR, Rodriguez DJ, Tran M, Feng Y, Viest NA, Dill GM (2002) Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiol* 129:1265-1275
- Bell MS, Tranel PJ, Riggins CW (2009) Glyphosate resistance in waterhemp: inheritance and EPSPS copy number. *Proc North Central Weed Sci Soc* 100.
<http://www.ncwss.org/proceed/2009/Abstracts/100.pdf>. Accessed June 25, 2014
[Abstract]
- Bell MA, Hager AG, Tranel PJ (2013) Multiple resistance to herbicides from four site-of-action groups in waterhemp (*Amaranthus tuberculatus*). *Weed Sci* 61:460-468
- Chatham LA, Riggins C, Owen MD, Tranel P (2012) Association of EPSPS gene amplification with glyphosate resistance in waterhemp. *Proc North Central Weed Sci Soc* 60.
<http://wssaabstracts.com/public/15/abstract-60.html>. Accessed June 25, 2014 [Abstract]
- Chatham LA, Wu C, Riggins CW, Hager AG, Young BG, Roskamp GK, Tranel PJ (2014) Association of EPSPS gene amplification with glyphosate resistance in Illinois waterhemp (*Amaranthus tuberculatus*) populations. *Weed Tech.* submitted
- Délye C, Duhoux A, Pernin F, Riggins CW, Tranel PJ (2014) Research methods in weed science: molecular mechanisms of herbicide resistance. *Weed Sci* 62: In press
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- Duke SO, Powles SB (2008) Glyphosate: a once-in-a-century herbicide. *Pest Manag Sci* 64:319-325

- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc Natl Acad Sci USA* 107:1029-1034
- Gaines TA, Shaner DL, Ward SM, Leach JE, Preston C, Westra P (2011) Mechanism of resistance of evolved glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*). *J Agric Food Chem* 59:5886-5889
- Gaines TA, Wright AA, Molin WT, Lorentz L, Riggins CW, Tranel PJ, Beffa R, Westra P, Powles SB (2013) Identification of genetic elements associated with EPSPS gene amplification. *PLoS one* 8(6): e65819
- Ge X, André d' Avignon D, Ackerman JJH, Sammons RD (2010) Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. *Pest Manag Sci* 66:345-348
- Hager AG, Wax LM, Stoller EW, Bollero GA (2002) Common waterhemp (*Amaranthus rudis*) interference in soybean. *Weed Sci* 50:607-610
- Heap IM (2014) International Survey of Herbicide Resistant Weeds. www.weedscience.org. Accessed May 15, 2014.
- Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Sci* 56:582-587
- Lorraine-Colwill DF, Powles SB, Hawkes TR, Hollinshead PH, Warner SAJ, Preston C, (2003) Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. *Pestic Biochem Physiol* 74:62-72

- Ma R, Kaundun SS, Tranel PJ, Riggins CW, McGinness DL, Hager AG, Hawkes T, McIndoes E, Riechers DE (2013) Distinct detoxification mechanisms confer resistance to mesotrione and atrazine in a population of waterhemp. *Plant Physiol* 163:363-377
- Nandula VK, Ray JD, Ribeiro DN, Pand Z, Reddy KN (2013) Glyphosate resistance in tall waterhemp (*Amaranthus tuberculatus*) from Mississippi is due to both altered target-site and nontarget-site mechanisms. *Weed Sci* 61:374-383
- Patton BP, Witt W, Martin JR (2012) Multiple resistance issues within Kentucky waterhemp populations. *Proc Weed Sci Soc Proc* 176. <http://wssaabstracts.com/public/9/abstract-176.html>. Accessed June 25, 2014 [Abstract]
- Powles SB (2010) Gene amplification delivers glyphosate-resistant weed evolution. *Proc Natl Acad Sci USA* 107:955-956
- R Development Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <http://www.R-project.org> Accessed March 26, 2014
- Ribeiro DN, Pan Z, Duke SO, Nandula VK, Baldwin BS, Shaw DR, Dayan FE (2014) Involvement of facultative apomixes in inheritance of *EPSPS* gene amplification in glyphosate-resistant *Amaranthus palmeri*. *Planta* 239:199-212
- Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW Scott RC Burgos NR (2012) EPSPS gene amplification in glyphosate-resistant Italian Ryegrass (*Lolium perenne* ssp. *Multiflorum*) from Arkansas, USA. *Pest Manag Sci* 68:1223-1230
- Schultz JL, Chatham LA, Riggins CW, Tranel PJ, Bradley KW (2014) Distribution of herbicide resistances and molecular mechanisms conferring resistance in Missouri waterhemp (*Amaranthus tuberculatus*) populations. *Weed Sci*. submitted

- Shaner DL, Lindenmeyer RB, Ostlie MH (2011) What have the mechanisms of resistance to glyphosate taught us?. *Pest Manag Sci* 68:3-9
- Steckel LE, Sprague CL (2004) Common waterhemp (*Amaranthus rudis*) interference in corn. *Weed Sci* 52:359-364
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30:2725-2729
- Tranel PJ, Trucco F (2009) 21st-century weed science: a call for *Amaranthus* genomics. Pages 53–81 in Stewart Jr. CN ed *Weedy and Invasive Plant Genomics*. Ames, IA: Blackwell
- Tranel PJ, Riggins CW, Bell MS, Hager AG (2010) Herbicide resistances in *Amaranthus tuberculatus*: a call for new options. *J Agric Food Chem* 59: 5808-5812
- Wakelin AM, Preston C (2006) A target-site mutation is present in a glyphosate-resistant *Lolium rigidum* population. *Weed Res* 46:432-440
- Wiersma S, Chisholm S, Godar A, Stahlman P, Leach J, Westra P (2012) Gene amplification of EPSP synthase in glyphosate resistant *Kochia scoparia*. *Proc Western Soc of Weed Sci* 128. <http://wssaabstracts.com/public/12/abstract-128.html>. Accessed June 25, 2014
- [Abstract]

3.7 Table and Figures

Table 3.1: Percentage of plants with elevated EPSPS copy number in the untreated control and glyphosate treatments for each population studied and the corresponding Chi-square goodness of fit p-values.

State	Percentage of plants with elevated EPSPS copy number		χ^2 p value
	Untreated	Glyphosate	
Illinois	32	85	<0.001
Kansas	31	89	<0.001
Missouri	75	80	0.386
Nebraska	45	86	<0.001

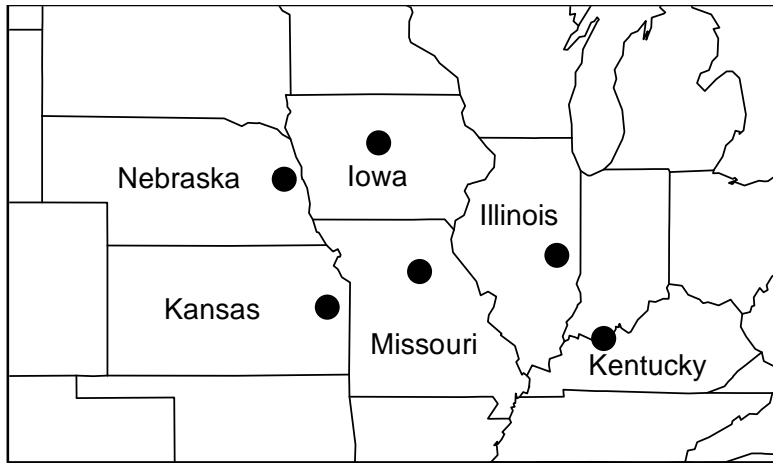


Figure 3.1: Map of the Midwest showing the locations of five glyphosate-resistant waterhemp populations investigated in this study plus an additional population (Iowa) included in a pilot study.

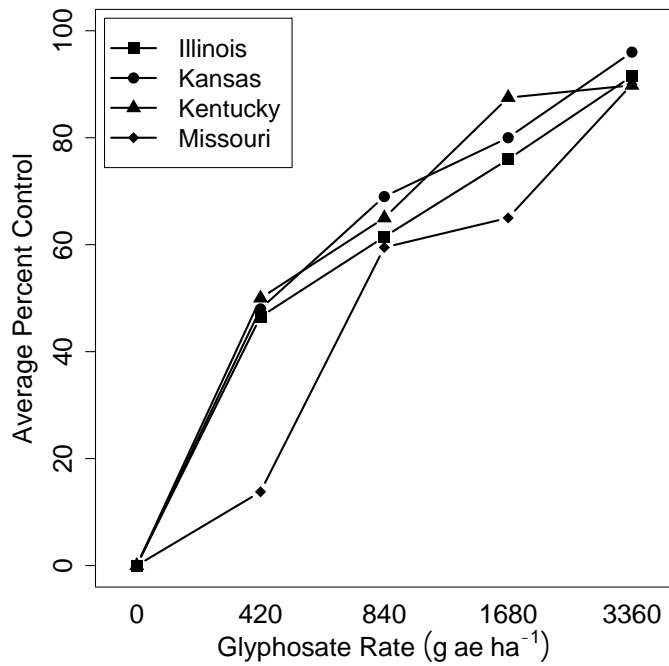


Figure 3.2: Results of the dose-response studies performed at each field location showing percent control with increasing glyphosate rate. Average percent control values are based on counts or visual observations of glyphosate activity and averaged for each rep at each location. (Data from Nebraska is not available)

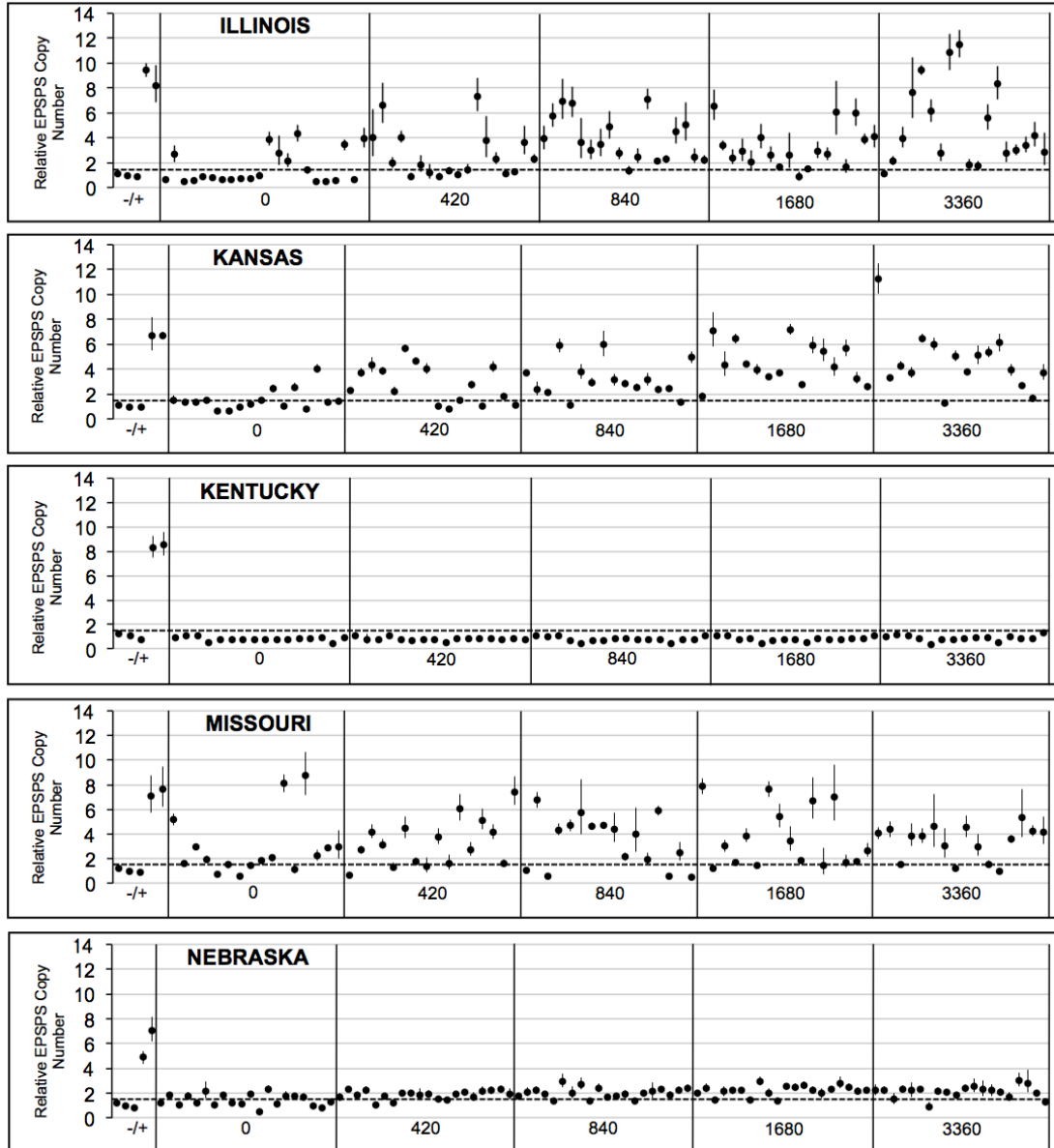


Figure 3.3: Raw copy number data showing relative EPSPS copy number of each surviving plant at each glyphosate rate (g ae ha^{-1}) in glyphosate-resistant populations from Illinois, Kansas, Kentucky, Missouri, and Nebraska. Relative EPSPS copy number data of three sensitive (-) and two resistant (+) controls are shown to the left of each graph. Solid vertical lines bound the copy number data from each glyphosate rate. Dashed lines represent the threshold EPSPS copy number value (1.5); samples with relative copy number above the threshold were considered to have EPSPS gene amplification.

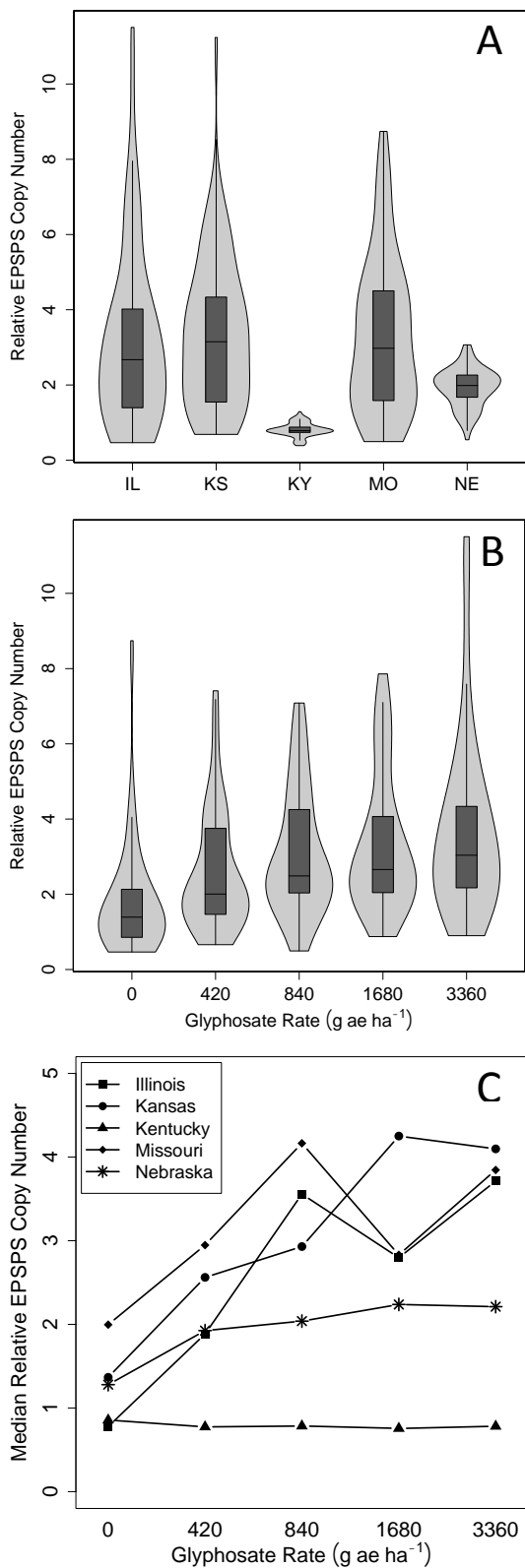


Figure 3.4: (A) Violin plots combining a standard boxplot with a kernel density plot to represent the distributions of relative EPSPS copy number combined for all treatments of each population, (B) Violin plots representing the distributions of the relative EPSPS copy number data at each glyphosate treatment level for the populations in which gene amplification was present (IL, KS, MO, NE), (C) Median relative EPSPS copy number of each population at each glyphosate rate.

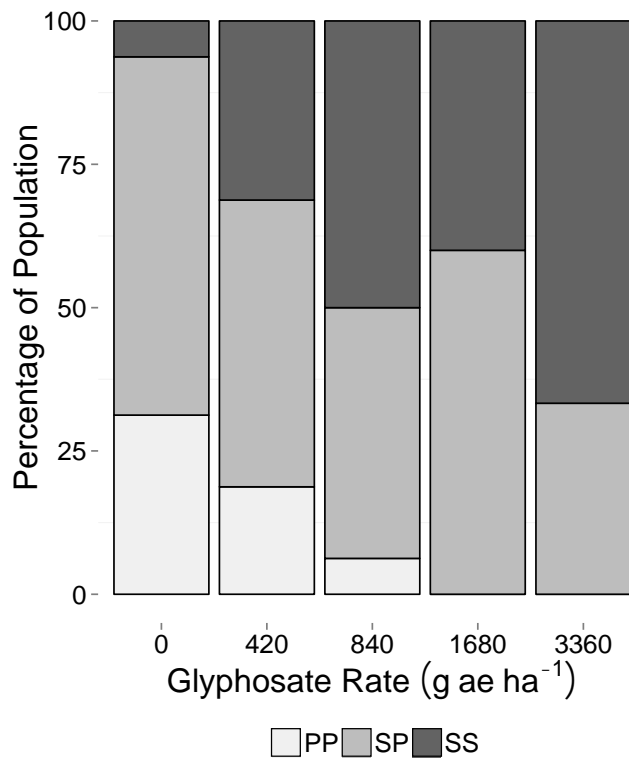


Figure 3.5: Percentage of survivors from the Kentucky population with the EPSPS Pro106Ser mutation at each glyphosate rate. PP individuals are homozygous with two proline alleles, SS individuals are homozygous with two serine alleles, and SP individuals are heterozygous.

CHAPTER 4

Examination of the EPSPS Pro106Ser Amino Acid Substitution in Glyphosate-Resistant Waterhemp (*Amaranthus tuberculatus*)

4.1 Abstract

Glyphosate is the most frequently used herbicide in the world, but the evolution and rise of glyphosate-resistant weeds threatens the sustainability of this herbicide as an effective tool for weed control. In the Midwest U.S., glyphosate-resistant waterhemp (*Amaranthus tuberculatus*) poses a significant economic threat. Three mechanisms of resistance have been identified in waterhemp: reduced uptake and/or translocation, EPSPS gene amplification, and target-site-based resistance via a Pro106Ser amino acid substitution. Two populations were recently identified with the Pro106Ser mutation, but questions remain regarding the extent to which the mutation confers resistance in these populations. Here we investigate the level of resistance endowed by the Pro106Ser substitution in waterhemp. A segregating F₂ population was created and progeny were screened for the mutation. A glyphosate dose response was then performed on individual genotypes, allowing for a comparison of genotypes as well as controlling for any differences in the background of the population. A 1:2:1 ratio of sensitive (PP) : heterozygous (SP) : homozygous (SS) was observed in the F₂ population. Resistance level increased with the addition of each allele containing the mutation. Heterozygous individuals were 2- to 3-fold more resistant than individuals in the same population without the mutation and individuals with two mutated alleles were 3- to 4-fold more resistant. A moderate level of glyphosate resistance, similar to that seen in other species with the mutation, appears to be endowed by the EPSPS Pro106Ser substitution in waterhemp.

4.2 Introduction

Glyphosate has been the most heavily used pesticide in the United States since 2001 (Grube et al. 2011). Its popularity is due in part to the efficiency with which it controls a broad spectrum of weeds, its accessibility and positive environmental profile, and with the advent of glyphosate-resistant (GR) crop technology, its applicability in a range of agronomic settings (Duke and Powles 2008). GR crop varieties are now available for soybeans, corn, cotton, canola, sugarbeets, and alfalfa. While the adoption of GR crops did allow for a decrease in the use of more persistent and toxic herbicides in favor of glyphosate, the excessive use and lack of diversity in herbicide sites-of-action used has favored the evolution of glyphosate-resistant weeds (Fernandes-Cornejo et al. 2014). Fourteen species have evolved resistance to glyphosate in the U.S. and 29 have evolved resistance worldwide (Heap 2014).

Glyphosate inhibits 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), an enzyme in the shikimic acid pathway that catalyzes the formation of 5-enolpyruvylshikimate-3-phosphate (EPSP) from phosphoenolpyruvate (PEP) and shikimate-3-phosphate (S3P). The shikimate pathway is responsible for the synthesis of aromatic amino acids as well as a number of important biochemical intermediates. Controversy remains regarding the actual mechanism of plant death, but most assume that the lack of aromatic amino acids leads to the prevention of protein synthesis, essentially starving the plant to death. Others suggest that deregulation of the shikimic acid pathway leads to a massive carbon flux away from other vital pathways (Duke and Powles 2008). Weeds have evolved several mechanisms of evading glyphosate, including reduced translocation (Lorraine-Colwill et al. 2003) via vacuolar sequestration (Ge et al. 2010), target-site mutations (Baerson et al. 2002), and amplification of the target-site gene (Gaines et al. 2010).

The first target-site mutation conferring glyphosate resistance was found in goosegrass (*Eleusine indica*) to be a single nucleotide polymorphism (SNP) resulting in a proline to serine amino acid substitution (Baerson et al. 2002). Subsequently SNPs in the EPSPS gene resulting in the substitution of proline for alanine (Yu et al. 2007), threonine (Ng et al. 2003), or leucine (Kaundun et al. 2011) have been identified. Six weed species to date have evolved mutations resulting in a substitution at Pro106, most of which are monocot species including *Lolium rigidum*, *Lolium multiflorum*, *Echinochloa colona*, *Eleusine indica*, and *Digitaria insularis*. Waterhemp (*Amaranthus tuberculatus*) was the first dicot species to evolve a mutation in the glyphosate target-site gene (Bell et al. 2013, Nandula et al. 2013). Studies in *E.coli* suggest that amino acid substitutions at the Pro106 site narrow the EPSPS binding pocket. Therefore, binding of the larger glyphosate molecule becomes unfavorable, while binding of the smaller substrate, PEP, remains mostly unaltered allowing EPSPS to retain its catalytic activity (Healy-Fried et al. 2007). Currently no studies examining the fitness of individuals with the Pro106Ser substitution have been reported in any weed species. However, enzyme kinetics of a Pro106Ser mutant created via site-directed mutagenesis of EPSPS in petunia showed a significant increase in $K_{m(app)}(PEP)$ from 5 μ M in the wild type enzyme to 44 μ M in the mutant (Padgett et al. 2001). These results indicating a decrease in EPSPS affinity for PEP suggest that a fitness penalty may be associated with SNPs resulting in a substitution at Pro106. Although subsequent research on goosegrass EPSPS indicated that the loss in substrate affinity was not as severe as originally suspected (Baerson et al. 2002).

With the incidence of stacked herbicide resistance traits and the resulting lack of effective herbicide chemistries available, waterhemp (*Amaranthus tuberculatus*) may be one of the most economically threatening of all GR weeds. In this one species, there is evidence for three

different glyphosate resistance mechanisms. EPSPS gene amplification was the first mechanism to be suspected (Tranel 2010, Shaner et al. 2011). Subsequently, a target-site mutation resulting in a proline to serine amino acid substitution (Bell et al. 2013, Nandula et al. 2013), and altered translocation and / or uptake (Nandula et al. 2013) were identified as potential resistance mechanisms. However, questions remain regarding the extent to which the Pro106Ser mutation contributes to glyphosate resistance. While Nandula et al. consistently found the mutation conferring the Pro106Ser substitution in all GR plants and never in GS plants, the presence of both altered target-site-based and non-target-site-based mechanisms makes it difficult to determine the respective contributions of each mechanism. Bell et al. (2013) identified the Pro106Ser mutation in plants that did not survive a 3360 g ae ha⁻¹ rate of glyphosate, suggesting that the mutation did not fully account for glyphosate resistance. The objective of this study was to determine the extent to which the Pro106Ser substitution confers glyphosate resistance in waterhemp.

4.3 Materials and Methods

4.3.1 Creation of a Segregating Line

Plants from the BCG population, previously identified as having the Pro106Ser substitution (Bell et al. 2013) were screened for glyphosate resistance. A GR BCG female was crossed with a glyphosate-sensitive (GS) MCR male from the original HPPD-resistant waterhemp population (Hausman et al. 2011). The progeny were screened with glyphosate to identify resistant individuals. Three surviving females and two surviving males were crossed to create three segregating pseudo-F₂ lines. Inflorescences from each female were collected and dried at room temperature. Seeds were threshed and stored at 4 C. A subset of seeds from each

female were stratified prior to germination to help break seed dormancy. First, seeds were soaked in 1:1 commercial bleach : deionized water solution for 10 minutes to surface sterilize the seeds and aid in breaking down the seed coat. Seeds were then rinsed twice with deionized water and then suspended in 0.1% agarose. Seeds were stored at 4 C for at least 6 weeks prior to use.

4.3.2 Screening for Pro106Ser Substitution

Newly emerging leaves approximately 1 to 2 cm in length were sampled from each plant used in the dose response (described below). DNA was extracted using a CTAB DNA extraction procedure essentially as described (Doyle and Doyle 1990). Samples were then screened for the Pro106Ser substitution using the dCAPS assay described previously (Chapter 2.3.3 Examination of Resistance Mechanisms). This assay detects the wild type; therefore it is possible that alternative amino acid substitutions (alanine and threonine) resulting from a single nucleotide polymorphism at the first nucleotide of the proline codon would be recognized.

4.3.3 Glyphosate Dose Response

Plant Culture

Seeds were germinated on filter paper moistened with deionized water and fitted in standard petri dish lids. Petri dishes were sealed with Parafilm (Pechiney Plastic Packaging, Menasha, WI 54952) to maintain a moist environment for germination and placed in a germination chamber (Conviron, 590 Berry St., Winnipeg, Manitoba, Canada R3H 0R9) set to cycle through 12-hr days at 35 C and 12-hr nights at 15 C. Petri dishes were removed 48 hours later, or when emerged hypocotyls had extended approximately 1 to 2 mm from the seed coat. Germinated seedlings were transplanted into 27.8 cm x 54.5 cm x 6.2 cm flats fitted with plastic

inserts containing 96, 2.9 cm x 4.0 cm x 5.7 cm individual cells (T.O. Plastics, 830 County Road 75, P.O. Box 37, Clearwater, MN 55320) filled with commercial potting mix (Sunshine Mix #1 / LC1, Sun Gro Horticulture, 770 Silver Street, Agawam, MA 01001). When plants reached 4 to 5 cm in height, a uniform subset of plants was chosen and each was transplanted into a 10.8 cm x 10.8 cm x 12.7 cm square plastic pot (Myers Industries Lawn & Garden Group, 15200 Madison Rd., P.O. Box 738, Middlefield, OH 44062) containing approximately 950 ml of growth medium consisting of commercial potting mix (Sunshine Mix #1 / LC1): soil : peat : sand. Plants were grown in a greenhouse room set to 16-hr days, with supplemental lighting from mercury halide and sodium vapor lamps to maintain a minimum of $800 \mu\text{mol m}^{-2} \text{sec}^{-1}$. Temperature was maintained at 24 to 26 C at night and 28 to 30 C during the day. Plants were fertilized with a complete slow-release fertilizer (Scott's Osmocote Classic 13-13-13, The Scotts Company, 1411 Scottslawn Rd., Marysville, OH 43041).

Experimental Design

Progeny from one of the F₁ females used to make the F₂ population were screened for glyphosate resistance and the response was compared to that of a GS control, ACR (Patzoldt et al. 2005), and a GR control, MO1 (Legleiter and Bradley 2008). These sensitive and resistant control populations were the same controls used in a previous study examining glyphosate resistance in the BCG population (Bell et al. 2013). The experiment was carried out as a completely randomized design, however the number of reps varied among the lines and genotypes studied. The first time the experiment was performed, plants were first genotyped and then placed into treatment groups prior to herbicide application. When the experiment was repeated, the number of F₂ plants used at each treatment was chosen based on the proportion of

each genotype present in the population. Enough plants were used to ensure that a minimum of four plants of each genotype (PP, SP, SS) were present at each treatment level. The number of plants of each genotype ranged from as few as four to as many as 21 plants per treatment level in the second run. In both experiments, there were at least four reps per treatment.

Herbicide Application and Data Collection

When plants were 10 to 15 cm in height, glyphosate (Roundup WeatherMax, Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, MO 63167) was applied at rates of 0, 52.5, 105, 210, 420, 1680, 3360, and 6720 g ae ha⁻¹ with 2.5% (v/v) ammonium sulfate (AMS) in a moving-nozzle cabinet spray chamber. All rates were used for the BCG population, but the ACR population was not sprayed with the two highest rates (3360, 6720 g ha⁻¹) and the MO1 population was not sprayed with the two lowest rates (0, 52.5 g ha⁻¹) The spray chamber was calibrated to deliver 187 L ha⁻¹ and was equipped with a 80015 even flat fan nozzle (TeeJet Technologies, P.O. Box 7900, Wheaton, IL 60187) positioned approximately 46 cm above the plant canopy. At 15 days after treatment (DAT) plants were rated visually on a scale of 0 to 100, where 0 represents no visible injury or stunting of growth and 100 indicates plant death. At 16 DAT aboveground biomass of plants was harvested and dried for at least five days at 65 C prior to recording dry weight.

Comparison of Sensitive Controls

Due to a limited availability of seed, ACR was used in place of MCR, the GS biotype used to create the F₂ population. To account for any potential difference in glyphosate response of sensitive controls, a small-scale glyphosate dose response was performed on ACR and MCR.

Plants were grown as described previously and treated when they reached 10 to 15 cm in height. Glyphosate (Roundup WeatherMax,) was applied at rates of 0, 52.5, 105, 210, 420, 840 and 1680 g ha⁻¹ with 2.5% (v/v) ammonium sulfate (AMS). Plants were rated visually at 16 DAT and dry weights were determined as described previously.

4.3.4 Dose Response Statistical Analysis

Preliminary analysis of the visual data and dry weight data suggested no significant treatment by run interaction ($\alpha=0.05$), therefore the data for each experimental run were pooled. Data were analyzed using a non-linear regression in the drc package (Analysis of dose-response curve data, Knezevic et al. 2007) in R (v3.0.1) (R Development Core Team 2013). A four-parameter log-logistic equation

$$y = c + \frac{d-c}{1+\exp\{b[\log(x)-\log(GR_{50})]\}} \quad [1]$$

(Seefeldt et al. 1995) was used to create the dose response model and determine the rate at which biomass was reduced by 50% compared to untreated controls. In this equation, y is the percent of control based on dry weight, x is the herbicide rate, c is the lower limit, d is the upper limit, and b is the slope of the curve.

4.4 Results and Discussion

An examination of the proportion of each genotype present in the F₂ population indicated that the Pro106Ser mutation likely follows a Mendelian pattern of inheritance. In the first run, 271 plants were screened for the Pro106Ser mutation and 20% were wild type, 48% were heterozygous and 31% were homozygous for the mutation. In the second run 21% of plants were wild type with two proline alleles, 50% were heterozygous and 29% were homozygous for the

mutation of 379 screened. Chi-squared goodness of fit tests were performed to determine if the data deviated significantly from the typical Mendelian 1:2:1 ratio. A modestly significant p value ($p=0.03$, $\alpha=0.05$) was observed for the data from the first run indicating that the observed percentages did deviate from the expected ratio. However, the p value for the data from the second run indicated no significant deviation ($p=0.15$, $\alpha=0.05$). Progeny from the two other F_1 females included in the cross to create a segregating F_2 line were screened for the Pro106Ser mutation. Results indicated that the observed percentages of each genotype did not significantly deviate from the expected ($p=0.90$, $p=0.25$, $\alpha=0.05$).

Analysis of the dry weight data from each run of the experiment indicated that there was no significant interaction between treatment and run, so the data were pooled for the generation of dose response curves and for estimation of GR_{50s} . The visual data results were very similar to the dry weight data, and the same patterns were evident in both. Visual inspection of the dose response curves (Figure 4.1) shows that individuals with at least one serine allele are resistant compared to the sensitive control (ACR) and the level of resistance also appears to increase with the addition of each serine allele. This suggests that the Pro106Ser substitution does confer some level of resistance to glyphosate. All genotypes from the BCG F_2 population were resistant compared to ACR, but less resistant than MO1. This suggests that an additional mechanism of resistance was present in the original BCG or MCR population from which the F_2s were made. This result echoes the conclusions made by Bell et al. (2013) who examined the original BCG population and concluded that the Pro106Ser mutation was not fully responsible for resistance.

GR_{50} values were generated for each line used in the dose response as well as R/S (resistant / sensitive) ratios using ACR or BCG F_2 PP individuals as the sensitive controls (Table 1). The R/S ratio for MO1 was 9.1, lower than previously observed in the original MO1

population (19-fold, Legleiter and Bradley 2008), but similar to the 6 to 13-fold value observed by Bell et al. (2013). The glyphosate dose required for a 50% reduction in biomass for ACR was $73 \pm 9 \text{ g ha}^{-1}$. Compared to ACR, plants heterozygous for the P106S substitution were 4.5-fold resistant, and homozygous plants were 6.6-fold resistant. These numbers are close to the 5-fold level of resistance found previously for this population (Bell et al. 2013) and in other waterhemp populations with the mutation (Nandula et al. 2013). However, compared to plants in the same F₂ population without the mutation, heterozygous plants were only 2.4-fold resistant and plants homozygous for the mutation were 3.5-fold resistant. This suggests that the Pro106Ser substitution may provide only a modest level of resistance and that some other supplementing mechanism of resistance is present in the population.

While ACR was the sensitive population originally compared to the BCG population (Bell et al. 2013) and the same used in this experiment, the difference in the observed level of resistance between ACR and the F₂ individuals without the Pro106Ser substitution warranted a comparison of ACR with the original sensitive parent used to make the F₂ population. Visual inspection of the dose response curve shows that the difference between the ACR and MCR curves is similar to the difference between the ACR and PP curves in the BCG F₂ dose response curve (Figure 4.1 and 4.2). The wild type PP individuals of the BCG F₂ line were 1.9-fold resistant compared to ACR, and the MCR population was 1.8-fold resistant compared to ACR (Table 1). Given how close the difference in resistance levels is in these two separate experiments, it is likely that the MCR population has some glyphosate tolerance that was inherited by the BCG F₂ population, explaining the difference in resistance level compared to the sensitive ACR control. However, these results differ from past examinations of this population. Bell et al. (2013) suspected that the Pro106Ser substitution was not singly responsible for

resistance in the BCG population, while here we found that the supplementing glyphosate tolerance was contributed by the MCR parent used to make the segregating F₂ population.

Given the results of both dose response experiments, the mutation conferring a Pro106Ser substitution provides an approximately 2 to 4-fold level of resistance. R/S ratios calculated from the visual data fell within this same range (data not shown). These results are similar to resistance levels observed in other GR species with the Pro106Ser substitution primarily responsible for resistance. A 2 to 4-fold glyphosate resistance level was observed in greenhouse studies of GR goosegrass (Baerson et al. 2002) and *E. coli* cells expressing the *EPSPS* gene from resistant goosegrass populations showed a greater than 3-fold level of resistance compared to cells expressing *EPSPS* from sensitive plants. A 2- to 5-fold level of resistance was observed in a *Lolium multiflorum* population with the Pro106Ser substitution (Jasieniuk et al. 2008) and a 2- to 3-fold resistance level was found in a *Lolium rigidum* population with a Pro106Thr substitution (Wakelin and Preston 2006). Many of the other populations identified with target-site resistance involving a substitution at Pro106 also have additional mechanisms of resistance making it difficult to determine the relative contribution of the target-site-based resistance mechanism (Sammons and Gaines 2014).

While waterhemp is the first dicot species to evolve target-site-based glyphosate resistance via a substitution at Pro106, the level of resistance seems to match closely to that observed previously in monocot species with the same mechanism of resistance. The presence of this mutation in waterhemp is particularly interesting given the widespread incidence of GR waterhemp populations with *EPSPS* gene amplification. While no studies have reported on the potential for a fitness penalty associated with the Pro106Ser substitution, the presence of multiple *EPSPS* copies may negate any disadvantages associated with acquiring the mutation.

EPSPS gene amplification may enable the evolution of the Pro106Ser substitution as well as other EPSPS mutations in populations that would have normally required a greater degree of selection. The combination of gene amplification and the Pro106Ser mutation in a single waterhemp population may prove interesting. The presence of the Pro106Ser substitution also presents the opportunity for the acquisition of additional mutations – specifically Thr101Ile. Jalaludin et al. (2013) reported the first case of an evolved double amino acid substitution in GR goosegrass containing the Pro106Ser mutation as well as the Thre101Ile mutation, the same mutations used to create the first GR maize varieties. Studies in *E.coli* suggest that the Thr101Ile mutation is unlikely to occur independently given the great increase in K_m for PEP that results (Funke et al. 2009). However with increased glyphosate selection, weed species with the Pro106Ser substitution may acquire the additional Thr101Ile substitution to confer a greater level of resistance. Without proper management techniques, the GR waterhemp situation in the Midwest may quickly intensify.

4.5 Literature Cited

- Baerson SR, Rodriguez DJ, Tran M, Feng Y, Biest NA, Dill GM (2002) Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiol* 129: 1265-1275
- Bell MA, Hager AG, Tranel PJ (2013) Multiple resistance to herbicides from four site-of-action groups in waterhemp (*Amaranthus tuberculatus*). *Weed Sci* 61:460-468
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- Duke SO, Powles SB (2008) Glyphosate: a once-in-a-century herbicide. *Pest Manag Sci* 64: 319-

- Fernandez-Cornejo J, Wechsler S, Livingston M, Mitchell L (2014) Genetically Engineered Crops in the United States, ERR-162. Washington, DC: U.S. Department of Agriculture, Economic Research Service
- Funke T, Yang Y, Han H, Healy-Fried M, Olesen S, Becker A, Schönbrunn (2009) Structural basis of glyphosate resistance resulting from the double mutation Thr⁹⁷ → Ile and Pro¹⁰¹ → Ser in 5-enolpyruvylshikimate-3-phosphate synthase from *Escherichia coli*. J Biol Chem 284:9854-9860
- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. Proc Natl Acad Sci U.S.A. 107:1029-1034
- Ge X, d'Avignon DA, Ackerman JJH, Sammons RD (2010) Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. Pest Manag Sci 66:345–348
- Grube AH, Donaldson D, Kiely T, Wu L (2011). Pesticides Industry Sales and Usage. Washington, DC: US EPA
- Hausman NE, Singh S, Tranel PJ, Riechers DE, Kaundun SS, Polge ND, Thomas DA, Hager AG (2011) Resistance to HPPD-inhibiting herbicides in a population of waterhemp (*Amaranthus tuberculatus*) from Illinois, United States. Pes Manag Sci 67:258-261
- Healy-Fried ML, Funke T, Priestman MA, Han H, Schonbrunn E (2007) Structural basis of glyphosate tolerance resulting from mutations of Pro101 in *Escherichia coli* 5-enolpyruvylshikimate-3-phosphate synthase. J Biol Chem 282: 32949-32955
- Heap IM (2014) International Survey of Herbicide Resistant Weeds. www.weedscience.org. Accessed June 27, 2014

- Jalaludin A, Han H, Yu Q, Powles S (2013) Evolution in action: a double amino acid substitution in the EPSPS gene endows high-level glyphosate resistance. Proc of Global Herbicide Resistance Challenge. <http://www.herbicideresistanceconference.com.au/Program>. Accessed July 1, 2014 [Abstract]
- Jasieniuk M, Ahmad R, Sherwood AM, Firestone JL, Perez-Jones A, Lanini WT, Mallory-Smith C, Stednick Z (2008) Glyphosate-resistant Italian ryegrass (*Lolium multiflorum*) in California: distribution, response to glyphosate, and molecular evidence for an altered target enzyme. *Weed Sci* 56:496-502
- Kaundun SS, Dale RP, Zelaya IA, Dinelli G, Marotti I, McIndoe E, Cairns A (2011) A novel P106L mutation in EPSPS and an unknown mechanism (s) act additively to confer resistance to glyphosate in a South African *Lolium rigidum* population. *J Agric Food Chem* 59: 3227-3233
- Knezevic SZ, Streibig JC, Ritz C (2007) Utilizing R software package for dose-response studies: the concept and data analysis. *Weed Technol* 21:840-848
- Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Sci* 56:582-587
- Lorraine-Colwill D, Powles S, Hawkes TR, Hollinshead P, Warner S, Preston C (2002) Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. *Pestic Biochem Physiol* 74: 62-72
- Nandula VK, Ray JD, Ribeiro DN, Pan Z, Reddy KN (2013) Glyphosate resistance in tall waterhemp (*Amaranthus tuberculatus*) from Mississippi is due to both altered target-site and nontarget-site mechanisms. *Weed Sci* 61: 374-383

- Ng C, Wickneswari R, Salmijah S, Teng Y, Ismail B (2003) Gene polymorphisms in glyphosate-resistant and-susceptible biotypes of *Eleusine indica* from Malaysia. *Weed Res* 43: 108-115
- Padgett SR, Re DB, Gasser CS, Eichholtz DA, Frazier RB, Hironaka CM, Levine EB, Shah DM, Fraley RT, Kishore GM (1991) Site-directed mutagenesis of a conserved region of the 5-enolpyruvylshikimate-3-phosphate synthase active site. *J Biol Chem* 266:22364-22369
- Patzoldt WL, Tranel PJ, Hager AG (2005) A waterhemp (*Amaranthus tuberculatus*) biotype with multiple resistance across three herbicide sites of action. *Weed Sci* 53:30-36
- R Development Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <http://www.R-project.org> Accessed March 26, 2014
- Sammons RD, Gaines TA (2014) Glyphosate resistance: state of knowledge. *Pest Manag Sci* 10.1002/ps.3743
- Seefeldt SS, Jensen JE, Fuerst EP (1995) Log-logistic analysis of herbicide dose-response relationships. *Weed Technol.* 9:218-227
- Shaner DL, Lindenmeyer RB, Ostlie MH (2012) What have the mechanisms of resistance to glyphosate taught us? *Pest Manag Sci* 68:3-9
- Tranel PJ, Riggins CW, Bell MS, Hager AG (2010) Herbicide resistances in *Amaranthus tuberculatus*: a call for new options. *J Agric Food Chem* 59: 5808-5812
- Wakelin AM, Preston C (2006) A target-site mutation is present in a glyphosate-resistant *Lolium rigidum* population. *Weed Res* 46:432-440

Yu Q, Cairns A, Powles S (2007) Glyphosate, paraquat and ACCase multiple herbicide resistance evolved in a *Lolium rigidum* biotype. *Planta* 225: 499-513

4.6 Table and Figures

Table 4.1: Glyphosate dose response results for the combined data of the F₂ dose response and the dose response for the comparison of sensitive controls where MO1 represents the resistant control, ACR is the sensitive control. Individuals in the BCG F₂ population without the Pro106Ser substitution are listed as PP, heterozygous individuals are listed as SP, and homozygous individuals in which both alleles carry the mutation are listed as SS. Results are based on dry weight of plants harvested 16 DAT. The glyphosate doses required to reduce biomass by 50% (GR₅₀) were estimated using a four parameter log-logistic model. R/S (resistant / sensitive) ratios are given using both ACR and PP individuals as the sensitive control.

Experiment	Population	GR ₅₀ (g ae ha ⁻¹)	R/S (ACR) ratio	R/S (BCG-PP) ratio
BCG Dose Response	BCG - PP	138 ± 22	1.9 ± 0.38	1.0
	BCG - SP	328 ± 29	4.5 ± 0.70	2.4 ± 0.43
	BCG - SS	481 ± 44	6.6 ± 1.04	3.5 ± 0.63
	MO1	660 ± 99	9.1 ± 6.97	4.8 ± 1.03
	ACR	73 ± 9	1.0	0.5 ± 0.11
Comparison of Sensitive Controls	MCR	90 ± 15	1.8 ± 0.37	
	ACR	50 ± 6	1.0	

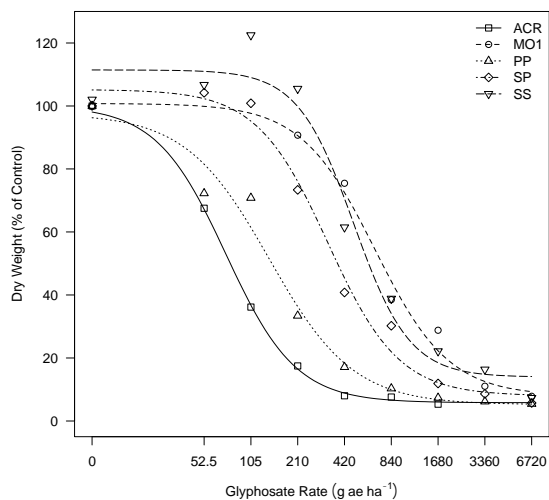


Figure 4.1: Dose response curve for the BCG F₂ glyphosate dose response created using the four parameter log-logistic model. Points represent mean dry weight of plants harvested 16 DAT converted to a percent of the untreated control.

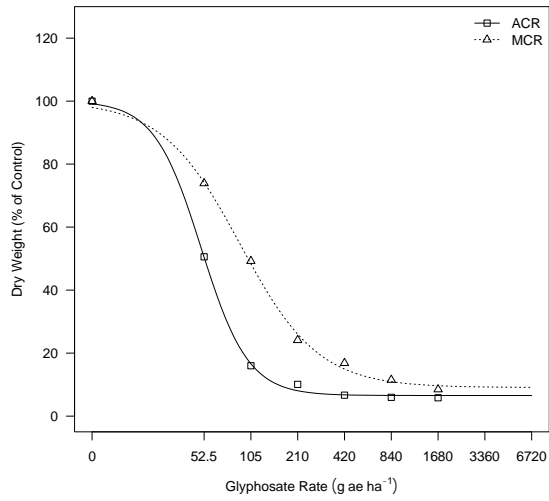


Figure 4.2: Dose response curve for the comparison of sensitive controls created using the four parameter log-logistic model. Points represent mean dry weight of plants harvested 16 DAT converted to a percent of the untreated control.

CHAPTER 5

Concluding Remarks

5.1 Summary and Conclusions

Glyphosate inhibits 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), an enzyme in the shikimic acid pathway responsible for the synthesis of aromatic amino acids in addition to a number of important biochemical intermediates. No other herbicides act on this same site of action (SOA), making glyphosate unique and highly valuable. Glyphosate binds tightly to the soil, making it relatively environmentally benign and, since only plants, fungi and some bacteria have EPSPS, glyphosate has low mammalian toxicity. Glyphosate is economical, especially since the U.S. patent has expired and generics are readily available, and it is highly effective. Systemic in nature, glyphosate phloem loads and is transported throughout the plant to sink tissues. There are also no known plants with pathways that can effectively degrade glyphosate, giving it a broad spectrum of control. The potency endowed by these traits allows glyphosate to be highly effective on perennials and other large plants allowing for flexibility in application. With the commercialization of glyphosate-resistant (GR) crops in 1996, glyphosate added crop selectivity to its already outstanding herbicide profile, making it the most widely used, most economically important herbicide worldwide (Duke and Powles 2008).

However, glyphosate may be too effective for its own good. The introduction of glyphosate, with its unique SOA presented an opportunity to increase diversity in weed control options and manage existing herbicide-resistant populations. Unfortunately the opposite happened upon the introduction of GR crops (Powles 2008). Many growers have abused the technology, relying solely on glyphosate for weed control, often using multiple applications per

year. Season after season, this overreliance on glyphosate and lack of diversity in management practices exerts tremendous selective pressure on weed species to evolve resistance.

Within two decades, waterhemp has grown from relative obscurity to become one of the most problematic weeds in the Midwest (Steckel 2007). It is extremely competitive with crop species and can dramatically reduce corn and soybean yields if not managed properly (Hager et al. 2002; Steckel and Sprague 2004). In 2005 the first GR waterhemp population was identified in a Missouri field that had been in continuous soybean production for at least six years and often treated with multiple applications of glyphosate (Legleiter and Bradley 2008). In subsequent years glyphosate resistance has spread throughout the Midwest; 15 states now have documented cases of GR waterhemp (Heap 2014). Despite the widespread occurrence of GR waterhemp throughout the Midwest, questions remain regarding the mechanism of resistance. Some have suggested that EPSPS gene amplification, the mechanism of glyphosate resistance in Palmer amaranth (*Amaranthus palmeri*), is also present in waterhemp, but the relationship between resistance and copy number has yet to be elucidated (Bell et al. 2009; Tranel et al. 2010; Shaner et al. 2012). Reduced translocation (Nandula et al. 2013) and a single nucleotide polymorphism in the EPSPS gene resulting in a Pro106Ser substitution in the mature EPSPS protein (Bell et al. 2013; Nandula et al. 2013) have been identified. Although in both populations containing the Pro106Ser substitution an additional mechanism of resistance was suspected making it difficult to thoroughly examine the extent to which the mutation confers resistance.

Results herein examine two potential target-site-based mechanisms of glyphosate resistance: 1) EPSPS gene amplification and 2) an EPSPS target-site mutation. Results presented in Chapter 2 showed that gene amplification was present in the majority of GR waterhemp populations sampled from Illinois, suggesting that elevated EPSPS copy number can be used as a

proxy for glyphosate resistance. The Pro106Ser substitution was identified in several of the sampled populations and some populations that were clearly resistant did not have either mechanism present. Chapter 3 looked deeper into the relationship between EPSPS gene amplification and glyphosate resistance. Results showed that when gene amplification was present, EPSPS copy number increased with increasing glyphosate dose. One of the populations studied did not have any evidence of gene amplification but instead had the Pro106Ser substitution, which was clearly responsible for resistance as complete selection for plants with the mutation was observed after the 1x rate. Chapters 2 and 3 both found evidence for the Pro106Ser substitution mentioned previously in the literature, so the research conducted in Chapter 4 was aimed at better understanding this mechanism and determining the level of resistance conferred. Results showed that, similar to monocot species with the same substitution, a moderate 2- to 4-fold level of resistance was conferred. Overall, EPSPS gene amplification appears to be the primary mechanism of glyphosate resistance in waterhemp, with some populations having a target-site Pro106Ser mutation conferring a more modest level of resistance. Substantial evidence for additional glyphosate resistance mechanisms was also obtained. In Chapter 2, resistant populations with neither gene amplification nor the Pro106Ser substitution were found. In Chapter 3, although each population had either gene amplification or the target-site mutation present, individual plants without gene amplification or the Pro106Ser substitution were found surviving at the highest glyphosate rates.

5.2 Implications and Future Directions

Waterhemp represents one of five weed species to have evolved EPSPS gene amplification. Others include Palmer amaranth (Gaines et al. 2010), the first identified, annual

ryegrass (*Lolium multiflorum*) (Salas et al. 2012), kochia (*Kochia scoparia*) (Wiersma 2012), and spiny amaranth (*Amaranthus spinosus*) (Nandula et al. 2014). Gene amplification was not anticipated as a possible mechanism of glyphosate resistance, but clearly the level of selection imparted on weed populations was strong enough to allow its evolution (Shaner et al. 2012). An obvious difference in the EPSPS gene duplication seen in these species is the variance in EPSPS copy number magnitude. Resistant Palmer amaranth populations have been identified with 40 to 100 relative copies in Georgia (Gaines et al. 2010), 20 to 60 in North Carolina (Chandi et al. 2012), and 33 to 59 in Mississippi (Ribeiro et al. 2014). Similar relative EPSPS copy number ranges have been seen in ryegrass, with 15 to 25 copies (Salas et al. 2012), and in spiny amaranth, with 33 to 37 copies (Nandula et al. 2014). In contrast, Kochia was found to have a relative copy number range of 3 to 9 (Wiersma 2012), and a population of Palmer amaranth from New Mexico had 2 to 8 relative EPSPS copies (Mohseni-Moghadam et al. 2013). These lower magnitude ranges are similar to those seen previously in waterhemp and in the research presented herein.

The drastic difference in relative copy number between these species raises a number of questions. Firstly, are 2 or 3 relative EPSPS copies sufficient to confer resistance? The suggestion that alternative mechanisms exist in the populations studied in this research make this question difficult to address. One of the conclusions from Chapter 3 was that higher glyphosate rates selected for plants with a higher relative EPSPS copy number. Gaines et al. (2010) also suggested that additional EPSPS copies confer a higher level of resistance. However, if only a few copies are required for resistance in waterhemp, and more copies result in a more resistant plant (as is suggested herein and in the literature), how are these species with only a few copies capable of withstanding such high rates of glyphosate? The GR Palmer amaranth from Georgia

in which a 40 to 100 relative genomic copy number range was observed (Gaines et al. 2010), was found to be 6- to 8-fold resistant compared to a GS population (Culpepper et al. 2006), and the GR spiny amaranth population with 33 to 37 relative EPSPS copies was 5-fold resistant relative to a sensitive control (Nandula et al. 2014). The original GR waterhemp populations identified as having an average of 4 relative EPSPS copies (Bell et al. 2009), was found to be 9- to 19-fold resistant compared to a GS population. Could there be more to the equation than just EPSPS amplification for waterhemp, kochia and the New Mexico Palmer amaranth population? Another puzzling piece of evidence comes from studies of rigid ryegrass and *Conyza* spp. that found 2 to 3-fold increases in EPSPS mRNA, but was not believed to be associated with resistance (Shaner et al. 2012). However, EPSPS mRNA expression and protein level were found to be positively correlated with EPSPS genomic copy number in kochia. This relationship has not yet been examined in waterhemp, but if the same holds true, the question of how some of these plants with only 2 to 3 copies survive is raised again. One explanation for these anomalies might be that a supplemental alternative resistance mechanism is acting in concert with gene amplification in these species. Future studies should examine gene amplification and resistance level conferred in these species side by side to better understand the manifestation of this mechanism across species. Future research efforts should also seek to identify and characterize any alternative mechanisms of resistance that may be present in waterhemp and then determine the effect on resistance level they may have in conjunction with gene amplification.

Furthermore, in species with tens of EPSPS copies, why did so many copies accumulate if only a few are needed, and by what evolutionary mechanism? The molecular mechanism of gene amplification may differ between the populations with high and low relative copy number ranges; examination of these mechanisms may prove insightful. Conversely if more copies are in

fact needed to survive higher levels of glyphosate, will EPSPS copy number in GR waterhemp populations increase to the levels observed in Palmer amaranth, annual ryegrass, and spiny amaranth? The results presented in Chapter 3 suggest that higher glyphosate rates select for plants with higher EPSPS copy numbers indicating that the higher the copy number, the higher the fitness of the individual. With more rounds of selection copy number may gradually increase in a population. We speculate that copy number magnitude may be a function of time and the rounds of selection experienced in Chapter 3. The Nebraska population which had lower overall copy number compared to the Illinois, Kansas and Missouri populations, was not suspected of glyphosate resistance until 2012, whereas the other populations were suspected resistant several years prior. The populations with higher copy number may have gone through more rounds of selection than the Nebraska population. This hypothesis that copy number magnitude is a product of the selection exerted however, is complicated by the possibility of alternative supplementing resistance mechanisms that may be present in the populations studied.

The other mechanism of glyphosate resistance studied in this research is the EPSPS target-site mutation resulting in a Pro106Ser amino acid substitution. Results from the dose response indicated that a 2- to 4-fold level of resistance is conferred. This is a relatively low level of resistance, and this mechanism was passed over during the creation of GR crops because the level of resistance was too low (Shaner et al. 2012). In a weed species under glyphosate selection however, the fitness advantage is realized. Weak mutations may be able to thrive in situations where herbicides are applied at reduced rates or to plants that are too large. Even if the plant is severely stunted or damaged, so long as the mechanism allows plants to make seed and reproduce, the mechanism is successful (Shaner et al. 2012). It is possible that the improper use of glyphosate enabled the evolution of the Pro106Ser substitution in waterhemp.

While waterhemp populations with this low level of resistance might be managed by higher glyphosate rates, there is the potential for a dangerous combination of this mechanism with EPSPS gene amplification. Additional studies would be needed to determine the level of resistance conferred by plants with amplified EPSPS copies containing the SNP encoding the Pro106Ser mutation. When gene amplification was first identified in Palmer amaranth (Gaines et al. 2010), Powles (2010) warned about the potential for the evolution of a highly resistant EPSPS given the buffering capacity endowed by multiple EPSPS copies. Whether the Pro106Ser substitution results in a significant fitness penalty has not yet been determined in weed species. However if one does exist, gene amplification may provide a handle for the evolution of the Pro106Ser substitution, in addition to other mutations in the EPSPS gene that might not have evolved independently. Another serious implication of the Pro106Ser mutation is the possibility that a more highly resistant double mutant may evolve in response to increased selection pressure. The first case of an evolved double amino acid substitution in EPSPS was recently reported in goosegrass (*Eleusine indica*) (Jalaludin et al. 2013). The double mutation (abbreviated TIPS) results in a Thr102Ile substitution in addition to the already evolved Pro106Ser substitution. This combination of substitutions results in a highly glyphosate-resistant EPSPS and was the method used to create the first GR maize varieties. Alone, the Thr102Ile substitution results in a large increase in K_m for the substrate phosphoenolpyruvate (PEP). In *E. coli*, Funke et al. (2009) observed a 9-fold increase in K_m , suggesting that the evolution of this mutation on its own would be highly unfavorable. However the observed K_m for the TIPS double mutant and the Pro106Ser mutation was only slightly reduced compared to wild type EPSPS. Funke et al. (2009) concluded that only the TIPS mutation would be sufficient to confer adequate glyphosate resistance while maintaining catalytic efficiency with PEP. While weed populations

were in fact able to develop resistance via a single mutation, these data suggest that the Thr101Ile substitution may only develop after the Pro106Ser given the drastic increase in K_m it may cause independently (Sammons and Gaines 2014). With increased selection however, weed populations with the Pro106Ser mutation may be forced to take this route and accumulate additional mutations that confer a higher level of resistance. Jalaludin et al. (2013) found that 78% of individuals in the double-mutant population survived the highest applied rate, 8640 g ae ha⁻¹, which is more than ten times the labeled 1x rate.

Without diligent management practices, the current GR waterhemp situation may quickly escalate. The commercialization of glufosinate-resistant crops and synthetic auxin-resistant crops presents an opportunity to increase the diversity of herbicide sites-of-action used. If glufosinate and the synthetic auxins are used appropriately with their resistant crop varieties, by stacking with multiple other herbicide modes of action and applying at the recommended rates, producers may be able to stave off or at least stall the evolution of weeds resistant to herbicides with these modes of action. Hopefully much has been learned from the loss of glyphosate as an effective method of weed control in some situations and hopefully the same mistakes will not be repeated with these new herbicide-resistant crop technologies.

5.3 Literature Cited

Bell MA, Hager AG, Tranel PJ (2013) Multiple resistance to herbicides from four site-of-action groups in waterhemp (*Amaranthus tuberculatus*). *Weed Sci* 61:460-468

- Bell MS, Tranel PJ, Riggins CW (2009) Glyphosate resistance in waterhemp: inheritance and EPSPS copy number. Proc North Central Weed Sci Soc 100.
<http://www.ncwss.org/proceed/2009/Abstracts/100.pdf>. Accessed June 25, 2014
[Abstract]
- Chandi A, Milla-Lewis SR, Giacomini D, Westra P, Preston C, Jordan DL, York AC, Burton JD, Whitaker JR (2012) Inheritance of evolved glyphosate resistance in a North Carolina Palmer amaranth (*Amaranthus palmeri*) biotype. Int J Agron doi:10.1155/2012/176108
- Culpepper AS, Grey TL, Vencill WK, Kichler JM, Webster TM, Brown SM, York AC, Davis JW, Hanna WW (2006) Glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*) confirmed in Georgia. Weed Sci 54:620–626
- Duke SO, Powles SB (2008) Glyphosate: a once-in-a-century herbicide. Pest Manag Sci 64: 319-325
- Funke T, Yang Y, Han H, Healy-Fried M, Olesen S, Becker A, Schönbrunn (2009) Structural basis of glyphosate resistance resulting from the double mutation Thr⁹⁷ → Ile and Pro¹⁰¹ → Ser in 5-enolpyruvylshikimate-3-phosphate synthase from *Escherichia coli*. J Biol Chem 284:9854-9860
- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. Proc Natl Acad Sci U.S.A. 107:1029-1034
- Hager AG, Wax LM, Stoller EW, Bollero GA (2002) Common waterhemp (*Amaranthus rudis*) interference in soybean. Weed Sci 50:607-610

Heap IM (2014) International Survey of Herbicide Resistant Weeds. www.weedscience.org.

Accessed June 27, 2014

Jalaludin A, Han H, Yu Q, Powles S (2013) Evolution in action: a double amino acid substitution in the EPSPS gene endows high-level glyphosate resistance. Proc of Global Herbicide Resistance Challenge. <http://www.herbicideresistanceconference.com.au/Program>.

Accessed July 1, 2014 [Abstract]

Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. Weed Sci 56:582-587

Mohseni-Moghadam M, Schroeder J, Ashigh J (2013) Mechanism of resistance and inheritance in glyphosate resistant Palmer amaranth (*Amaranthus palmeri*) populations from New Mexico, USA. Weed Sci 61:517-523

Nandula VK, Ray JD, Ribeiro DN, Pan Z, Reddy KN (2013) Glyphosate resistance in tall waterhemp (*Amaranthus tuberculatus*) from Mississippi is due to both altered target-site and nontarget-site mechanisms. Weed Sci 61: 374-383

Nandula VK, Wright AA, Bond JA, Ray JD, Eubank TW, Molin WT (2014) EPSPS amplification in glyphosate-resistant spiny amaranth (*Amaranthus spinosus*): a case of gene transfer via interspecific hybridization from glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*). Pest Manag Sci 10.1002/ps.3754

Powles SB (2008) Evolved glyphosate-resistant weeds around the world: lessons to be learnt. Pest Manag Sci 64:360-365

Powles SB (2010) Gene amplification delivers glyphosate-resistant weed evolution. Proc Natl Acad Sci USA 107:955-956

- Ribeiro DN, Pan Z, Duke SO, Nandula VK, Baldwin BS, Shaw DR, Dayan FE (2014)
Involvement of facultative apomixes in inheritance of *EPSPS* gene amplification in
glyphosate-resistant *Amaranthus palmeri*. *Planta* 239:199-212
- Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW Scott RC Burgos NR (2012) EPSPS gene
amplification in glyphosate-resistant Italian ryegrass (*Lolium perenne* ssp. *Multiflorum*)
from Arkansas, USA. *Pest Manag Sci* 68:1223-1230
- Sammons RD, Gaines TA (2014) Glyphosate resistance: state of knowledge. *Pest Manag Sci*
10.1002/ps.3743
- Shaner DL, Lindenmeyer RB, Ostlie MH (2012) What have the mechanisms of resistance to
glyphosate taught us? *Pest Manag Sci* 68:3-9
- Steckel LE (2007) The dioecious *Amaranthus* spp.: here to stay. *Weed Technol* 21: 567-570
- Steckel LE, Sprague CL (2004) Common waterhemp (*Amaranthus rudis*) interference in corn.
Weed Sci 52:359-364
- Tranel PJ, Riggins CW, Bell MS, Hager AG (2010) Herbicide resistances in *Amaranthus*
tuberculatus: a call for new options. *J Agric Food Chem* 59: 5808-5812
- Wiersma A (2012) Regional Whole Plant and Molecular Response of *Kochia scoparia* to
Glyphosate. M.S. thesis. Fort Collins, CO: Colorado State University. 64 p

APPENDIX A

A User's Guide to High-Throughput Screening for Herbicide Resistance in Waterhemp

A.1 Cone-tainers

In order to achieve maximum screening efficiency, single waterhemp plants were grown in plastic cells 21 cm deep with a 3.8 cm top diameter, hereon referred to as cone-tainers (Ray Leach "Cone-tainer"TM single cell system; Stuewe and Sons Inc., 31933 Rolland Dr., Tangen, OR 97389; <https://www.stuewe.com/products/rayleach.php>). This cone-tainer size is referred to as the SC10 super; 98 of these will fit in a 7 cone x 14 cone arrangement in an RL98 rack.

Cone-tainers come in three plastic varieties: black recycled (SC10R), white low-density (SC10L), and UV-stabilized (SC10U). In the early stages of waterhemp screening the white low-density cone-tainers were used. Following a round of screening, the containers were washed with a large test tube brush and dipped in an oxidate solution. The oxidate solution must be requested; greenhouse staff will set up a sink containing the solution. This process is rather time consuming and inefficient. When the number of plants to be screened increased, the black recycled cone-tainers were used instead and these were disposed of after each use. The RL98 racks are still reused.

Cone-tainers and racks are provided by the Plant Care Facility (PCF) and are stocked in the PCF storeroom. One box of cone-tainers contains 1000 cells, enough to fill approximately ten racks full of cones. Given the high-throughput experiments often performed by members of the Tranel lab, one should notify the greenhouse staff in advance regarding the number of cone-tainers that will be needed. An order should be placed with the greenhouse staff at least 1 to 2 weeks prior to the start of the experiment. In an emergency, if one needs to use all or the

majority of the cones available in the storeroom, the greenhouse staff should be notified as soon as possible so they can place an order to restock for other greenhouse users. Racks are also provided by the greenhouse staff.

A.2 Soil

The media used to fill cone-tainers is a soil mix provided by the PCF called Weed Lite Mix. Weed Lite Mix is a 3:1:1:1 mix of commercial potting mix (Sunshine Mix #1 / LC1, Sun Gro Horticulture, 770 Silver Street, Agawam, MA 01001) : soil : peat : torpedo sand. Soil should be ordered at least a week prior to the start of an experiment. Soil will be delivered to the 1700 hallway alcove. It can be used directly from the cart if used within a few days. However, the number of soil carts available is limited; they should not be used for long-term soil storage. Soil should be transferred to the two white soil tubs labeled Tranel Lab in the 1700 alcove. These each hold half of a cubic yard of soil. If requested during ordering, and there is space available, the PCF staff will often transfer the soil themselves to the white bins.

The volume of each cone-tainer is 10 cubic inches. A large cart of soil is delivered in a 1 cubic yard cart and a small cart is delivered in a half cubic yard cart. A half cubic yard of soil should fill about 24 racks of cone-tainers. However, this depends on how much the soil is compacted during filling and how full the cart is filled when the soil order is processed. Typically a half cart of soil will fill between 17 and 20 cone-tainer racks. Before ordering soil the experimenter should determine how many racks will be needed; making a layout is often helpful.

A.3 Filling Cone-tainers

While filling cone-tainers may seem like a fairly straightforward task, multiple rounds of trial and error have resulted in a detailed procedure for filling cone-tainers. In an effort to avoid excess soil on the greenhouse room floor, jumbo size cotton balls can be placed in the bottom of each cone prior to filling with soil. In the past, 400-count bags of White Cloud Jumbo Size Cotton Balls have been purchased from Walmart online. These can be ordered by Patty Sarver in AW-101 Turner Hall. One bag will fill about four racks of cone-tainers. To do this, a rack should first be completely filled with cones. Cotton balls should be placed in each cone and then pushed to the bottom using a PCV pipe of sufficient length. Skipping this step may require more frequent sweeping and cleaning under the benches.

After cotton balls have been placed in the bottom of each cone, one can begin to fill the cones with soil. The entire rack should be placed in a white soil tub on a flattened area, so soil can be easily scooped into the cones without spilling an excessive amount on the floor. When all cones are full of soil, the filler should move cone by cone pushing one or two fingers into the soil to remove air pockets. Cones should then be topped off with an additional layer of soil, creating soil filled cones that are flush with the top of the container. Soil will compact upon watering, leaving about a ¼” of space between the soil surface and the top of the container.

At this point, a clean, empty rack should be placed on the cart that will be used to transport cones back to the greenhouse room. Cones can now be removed from the rack used for filling and placed into the clean rack. While tedious, this step ensures that no soil will be trapped between cones. Skipping this step and / or not using cotton balls in the bottom of each cone will result in an excessive amount of soil falling on the floor of the greenhouse room when the cones are placed under the overhead misters. Depending on desired planting date, adequate time should

be reserved for filling cone-tainers. Filled cones should be left under misters for at least two misting sessions (~ 24 hours) prior to planting in order to wet the soil.

A.4 Germination

Seeds should be stratified and stored at 4 C for at least 6 weeks prior to germination to sterilize, help break seed dormancy, and promote uniform germination. To do this, seeds should first be soaked in a 1:1 commercial bleach : deionized water solution with occasional vortexing. After removing the bleach solution, seeds should then be rinsed in deionized water twice for 10 minutes to remove any residual bleach. Prior to germination for the experiment, it is often helpful to determine the germination rate by starting a small subset of seeds, especially if the experimenter is working with a population with limited or valuable seed.

Materials needed for germination include: sterile petri dishes, filter paper, ddH₂O, metal spatulas, parafilm, scissors, forceps, and gloves. Petri dishes should be labeled with the seed type, your name, and date. Two filter papers should be placed in the lid of each petri dish, and moistened with ddH₂O, about 4 to 5ml is an appropriate amount. Using forceps or a metal spatula, the paper should be pressed down and smoothed to remove any air bubbles. Desired seeds should be placed on the wetted filter paper and spread around (Figure A.1a). Spreading the seeds will make picking up a single seed easier during transplanting. The bottom of the petri dish can then be used to cover the seeds. A strip of Parafilm (approximately 1 in) should be cut and stretched around the petri dish to seal in moisture (Figure A.1b). Plates should be placed in the germination chamber in room N315, Turner Hall (Figure A.2). If the experiment is not associated with the weed ecology group, approval to use the germination chamber should be obtained from Dr. Adam Davis prior to use. If the germination chamber is not available, seeds

have been shown to germinate with varied success in petri dishes placed on benches in the greenhouse room. Using the germination chamber, settings for waterhemp should be 35 C day, 15 C night with 12-hour photoperiods (Figure A.3). A tray of water should always be placed in the bottom of the germination chamber to increase relative humidity. After 48 hours, germinated seeds are ready for transplanting.

A.5 Transplanting

Plastic labels, a permanent marker, metal spatulas, forceps, and germinated seeds should be taken to the greenhouse and pre-moistened, soil-filled cone-tainers should be transported to one of the workbenches in the main hall of the greenhouse. Before planting, labels with the appropriate seed name should be placed in the cones; making a layout first is often helpful (Figure A.4). Using a metal spatula, a small “divot” should be made in each cone-tainer to be planted in (Figure A.5). Forceps are then used to gently select a germinated seed and place it in the soil divot. Seeds should be held very gently with forceps by the emerging hypocotyl. However, no squeezing of the forceps is often required since the moist germinated seedlings will often just stick to the forceps. While uniform seedlings should be chosen to help ensure uniformity of mature plants, an appropriate seed for transplanting should have a hypocotyl only 2 to 3mm in length. Seedlings with longer hypocotyls are more prone to damage during transplanting. After seeds have been placed in all divots, the metal spatula should be used to gently close the soil in around the seed so that only a very light layer of soil covers the seedling. Once an entire section of a specific population or seed line has been planted, the label should be turned around to indicate that those cones have already been planted.

Because waterhemp often does not grow uniformly, extra seedlings should be transplanted. For the Illinois Statewide Survey Study, approximately two times as many germinated seedlings as needed were planted. However, the number of extra plants needed will depend on the population – if it is a population that has been used in the past, the experimenter should know about how many extra plants might be needed. The number of extra plants that might be needed should also be taken into consideration during germination. Appropriate positive and negative controls should also be planted at this time and with sufficient extra plants to achieve uniform control populations. Suggestions for positive and negative controls to be used for screening with various herbicides can be found in Table 1.

If all populations have germinated well, petri dishes can be disposed of by either autoclaving the fresh germinated seedlings or by throwing away the petri dishes after the seedlings have dried up and died. However, it is not uncommon for some populations to require additional germination time past the initial 48 hours. These petri dishes should be re-moistened, wrapped in Parafilm again, and returned to the germination chamber for another 24 hours.

A.6 Plant Care

General information about growing waterhemp and performing experiments in the greenhouse can be found in Appendix A: A Practical Grower's Guide to Experimentation on Waterhemp of Mike Bell's Master's thesis (available: <http://hdl.handle.net/2142/18456>).

A.6.1 Fertilizing

Care should be taken to ensure that plants grown in cone-tainers receive adequate nutrients. Plants should be dark green in color with large leaves. Without adequate nutrients,

plants will become yellow-green in color, internode length will increase, leaves will tend to be smaller than average, and plants will be clearly stressed and may flower early (Figure A.6). Plants should be fertilized when emerged seedlings start to grow true leaves, when plants are close to spray height (4 to 5 cm) and again approximately one week after herbicide application. Each cone should receive about 10 to 15 slow-release pellets each time it is fertilized. In the past, slow release pellets have been mixed in with the soil prior to filling the cone-tainers. If this is done, plants will still likely need additional fertilizer at some point. A complete slow-release fertilizer (Scott's Osmocote Classic 13-13-13, The Scotts Company, 1411 Scottslawn Rd., Marysville, OH 43041) has been used in the past. In all the cone-tainer experiments performed, there has never been a case of too much fertilizer applied.

In case the plants do begin to look unhealthy, they should be fertilized as soon as possible with liquid fertilizer (20-20-20). This will be the fastest way to help plants recover. Fertilizer can be applied directly by the user with the Dosatron injector located in the 1600 hallway of the PCF (Figure A.7). Alternatively, one can ask the greenhouse staff for assistance in fertilizing the plants.

A.6.2 Overhead Misting

After planting, cone-tainer racks should be placed back on the benches with overhead misters (Figure A.8a). The misters should be set to mist for 30 minutes, twice daily. As the plants grow, the settings may need to be changed to increase the amount of time the misters are on to 45 minutes. The appropriate misting time will allow the soil to dry between misting, but not allow wilting. The user can change the misting times, but one should contact a member of the greenhouse staff before using the mist controller for the first time (Figure A.8b). Prior to

spraying, it is important to move racks out from under the misters to ensure that plants are not wet for herbicide application. For this reason it is often helpful to know what times the misters turn on for each bench.

A.7 Spraying

Plants should be sprayed when they reach 4 to 5 cm in height, or when they have 4 to 5 true leaves (Figure A.9). Prior to spraying, a fair amount of time should be spent organizing the plants using the extra plants grown for each population. Plants should be arranged so that each population to be sprayed is relatively uniform in size. Populations are often grouped together with some space between populations to help prevent competition and shading. Once the populations have been made uniform, extra plants should be disposed of as soon as possible to make room for other experiments, especially given the large number of extra plants that are often needed. Given the variability between waterhemp populations, it is not uncommon to need to spray on multiple days in order to spray each population at the correct height and developmental stage. If spraying on several days, a tag should be put in each rack indicating the date it was sprayed. This helps the experimenter to know when each rack needs to be rated. After spraying, care should be taken to ensure that plants will not be immediately misted. The user may need to temporarily store the sprayed racks on other benches without misters. Alternatively, if no other cone-tainers are on the misting bench, the mist can be turned off temporarily. The user can do this after learning how to operate the spray box from a member of the greenhouse staff.

Additional information on mixing herbicides and spraying can be found in Appendix C: Herbicides and Liquid Fertilizer Mixing Instructions of Mike Bell's Master's thesis (available: <http://hdl.handle.net/2142/18456>)

A.8 Rating

Plants are usually rated 14 days after spraying. The scale in Table 2 gives a general idea of how plants should be rated. Example plants can be found in Figure A.10. Unfortunately, due to the variety of responses to different herbicides, this scale does not adequately describe all possible intermediate phenotypes. For glyphosate resistance screening there is often a wide range of intermediate plants that are not quite sensitive but not nearly as strong as the positive control or the 0x controls. An example of sensitive and resistant populations can be seen in Figure y. Due to the wide range of responses, it is very helpful to take pictures of each rack after rating. Populations can then be compared visually after all rounds have been completed. Pictures may also be useful when multiple experimenters have rated the plants, or when comparing to a previous lab member's results (Figure A.11). Usually pictures are taken as an overhead shot of each rack. However, if plants are especially tall, side shots may be needed in order to see each individual. After rating and taking pictures, any tissue samples that might be needed from resistant or intermediate individuals should be taken.

A.9 Disposal and Cleanup

Once the experiment is over, plants should be disposed of quickly to free space for other experiments. Plants should be clipped at the base and put in an autoclave bag. Later this plant material should be autoclaved and the bag placed inside an opaque black bag before being thrown away. Alternatively, if asked, the greenhouse staff will autoclave and dispose of the material. After all aboveground plant material has been removed, soil should be dumped from each cone-tainer into the "soil and plant material only" trash can provided by the greenhouse. The cones can be thrown in the regular "trash only" cans provided. Racks should be taken to the

hallway and placed in the dirty pot collection bins. The greenhouse staff will wash these and then restock them in the supplies room.

A.10 Out-of-Season Experiments

Waterhemp grows much faster in the summer months when the days are longer, than in the winter when days are shorter. It also tends to flower earlier in the winter months. In general, cone-tainer experiments should be avoided during times of the year when day length is the shortest (late November through early February). During the summer, the time required to reach spray height is much shorter, and the plants are usually much larger by two weeks after treatment. During the fall and spring, when days are somewhat shorter, it may take more time to reach spray height and the plants may not be as tall by rating time. During these times, one may notice that the plants lean towards the windows, the source of the strongest natural light. To avoid leaning plants, the experimenter should rotate the racks every several days.

A.11 Literature Cited

- Hausman NE, Singh S, Tranel PJ, Riechers DE, Kaundun SS, Polge ND, Thomas DA, Hager AG (2011) Resistance to HPPD-inhibiting herbicides in a population of waterhemp (*Amaranthus tuberculatus*) from Illinois, United States. *Pes Manag Sci* 67:258-261
- Legleiter TR, Bradley KW (2008) Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Sci* 56: 582-587
- Patzoldt WL, Tranel PJ, Hager AG (2005) A waterhemp (*Amaranthus tuberculatus*) biotype with multiple resistance across three herbicide sites of action. *Weed Sci.* 53:30–36

A.12 Tables and Figures

Table A.1

Suggested sensitive and resistant controls for various herbicide resistance screening in waterhemp.

Herbicide SOA used for Screening	Sensitive Controls	Resistant Controls
ALS (e.g. Pursuit, Classic)	WCS ¹	ACR ²
EPSPS (e.g. WeatherMax, PowerMax)	WCS ¹ , WT ³ , ACR ²	MO1 ⁴
HPPD (e.g. Callisto, Laudis)	WCS ¹ , WT ³ , ACR ²	MCR ⁵
PPO (e.g. Cobra, Flexstar)	WCS ¹	ACR ²
PSII (e.g. Aatrex, Sencor)	WCS ¹ , MO1 ⁴	ACR ² , MCR ⁵

¹ WCS: Wayne County Sensitive (Patzoldt et al. 2005)

² ACR: Adams County Resistant (Patzoldt et al. 2005)

³ WT: Wild Type – an assortment of seed from populations collected throughout the state in 2003

⁴ MO1: Missouri 1 (Legleiter and Bradley 2008)

⁵ MCR: Mclean County Resistant (Hausman et al. 2011) – also, NH1 or NH2 (Nick Hausman 1 and 2): progeny of MCR x MCR cross

Table A.2

Herbicide resistance rating scale used for glyphosate resistance screening in cone-tainers.

R / I / S Rating	Numeric	Criteria
Sensitive (S)	0	Plant is completely dry, no live tissue left, no chance of regrowth
Sensitive (S)	1	Plant has some live stem tissue but meristem and leaves are dead, no chance of regrowth
Intermediate (I)	2	Plants are no longer erect, growth is severely stunted, obvious chlorosis or necrosis, plants may or may not continue growing
Intermediate (I)	3	Growth is severely stunted but plants are upright and erect, some evidence of chlorosis or necrosis
Resistant (R)	4	Plants are smaller in size than 0x controls, but otherwise appear unharmed, plants will continue growing
Resistant (R)	5	Plants are approximately the same size as the 0x controls, plants will continue growing

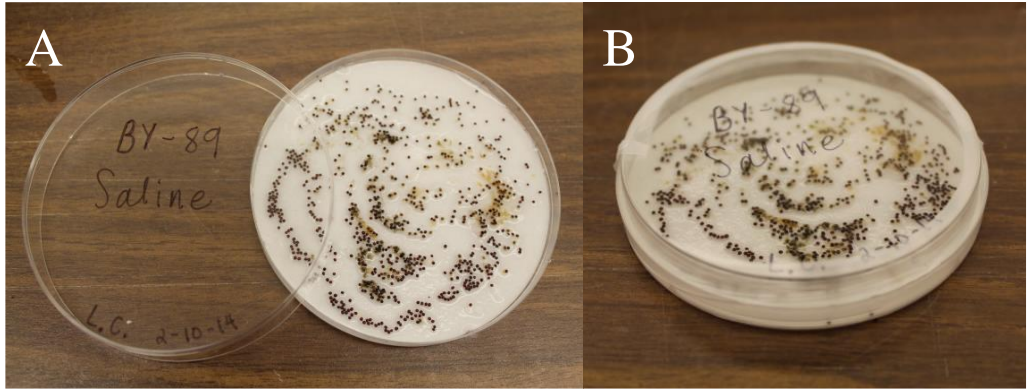


Figure A.1: A) Petri dish showing how seeds should be germinated, B) Petri dish wrapped in Parafilm and ready to be put in germination chamber. For this particular experiment, a large number of seeds were used because the germination rate was unknown and seed was not particularly valuable.



Figure A.2: Germination chamber located in room 315 Turner Hall. Water should be kept in the pan in the bottom of the chamber to increase humidity.



Figure A.3: Germination chamber screen showing correct settings for germinating waterhemp.

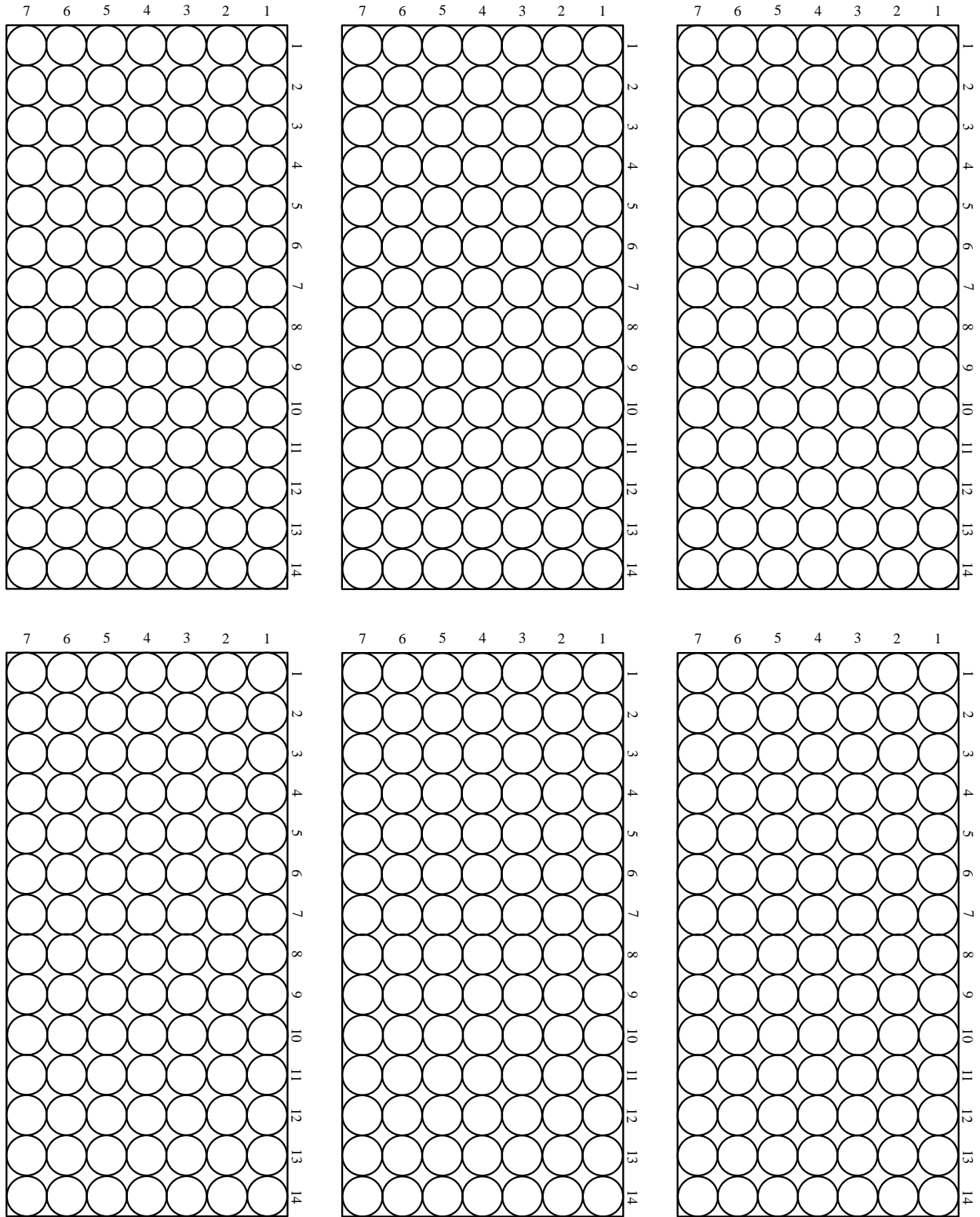


Figure A.4: Layout for planning cone-tainer experiments. This may be especially helpful for large-scale experiments.



Figure A.5: Cone-tainers ready for transplanting. Divots should be made in the soil of each cone.

A seedling should then be placed in each and covered gently with a thin layer of soil.

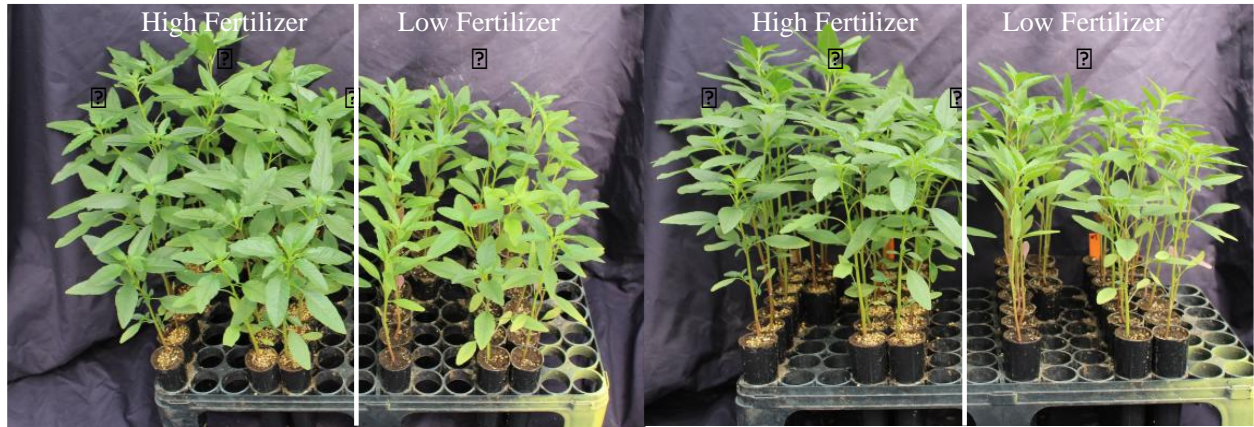


Figure A.6: Results of an experiment growing waterhemp in cone-tainers with high (fertilized generously three times) and low amounts (fertilized once) of a complete slow-release fertilizer (13-13-13). Results show the importance of applying enough fertilizer.



Figure A.7: The Dosatron cart used for liquid fertilizer application. Users should ask for help from a member of the Plant Care Facility staff prior to operating for the first time. The 20-20-20 concentrate option is what is normally used for waterhemp.

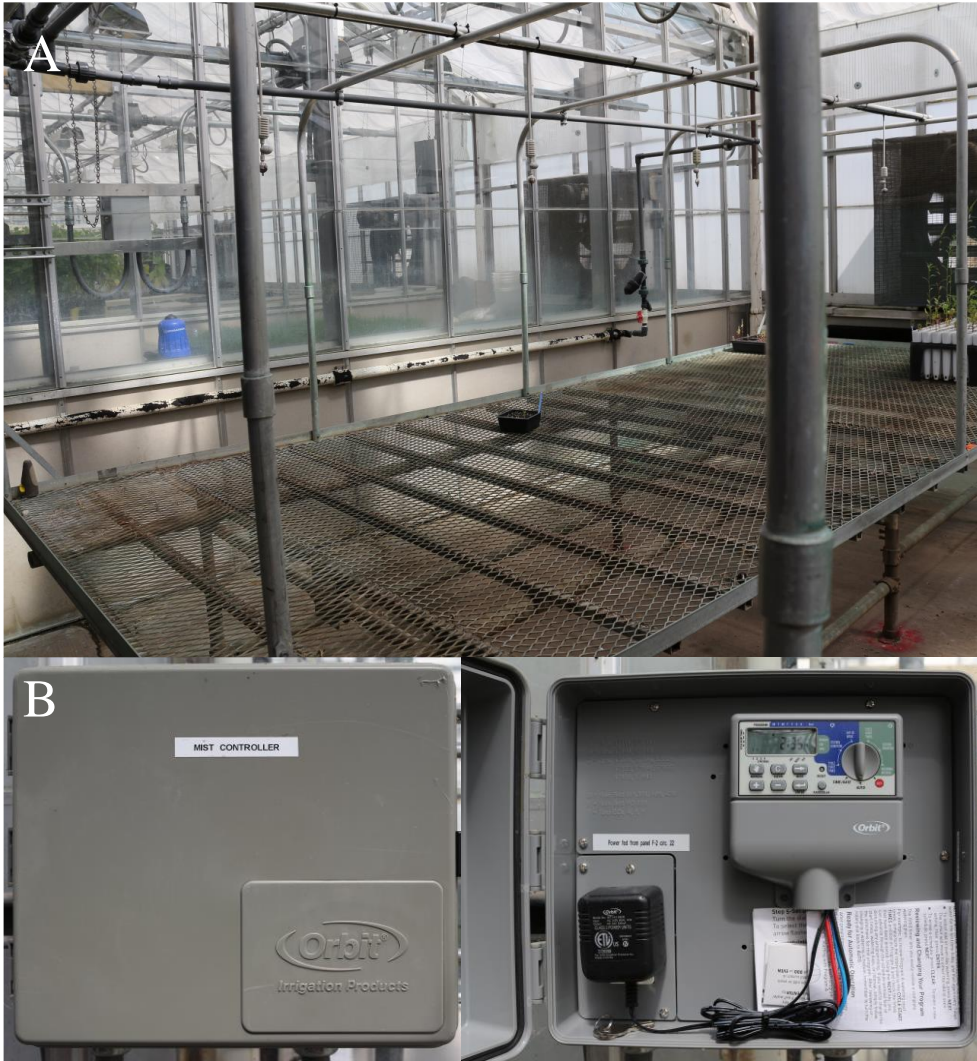


Figure A.8: The misting system used for watering cone-tainer experiments in room 1707 PCF.

A) A misting bench with overhead irrigation. B) The mist controller located on the west side of 1707 next to the faucet.



Figure A.9: A relatively uniform population of waterhemp, ready to be sprayed. Most of the plants in this photo have four true leaves, but some are starting to make their fourth and some are already making a fifth. Some variation is to be expected.



Figure A.10: Rating scale corresponding to Table 1. Plants in this picture were treated with glyphosate two weeks prior. Ratings of 0 and 1 are considered sensitive, 2 and 3 are considered intermediate, and 4 and 5 are resistant. While this scale provides a starting point for rating plants, the variability in injury symptoms for different herbicides and the wide range of intermediate responses (especially with glyphosate) should be taken into account when determining a rating scale.



Figure A.11: An example rack of plants from the Illinois Statewide Survey Study. Pictures were taken two weeks after treatment with glyphosate. A sensitive control population of seven plants is on the far left, followed by a sensitive population and two resistant populations, each of fourteen plants, and a resistant control population of seven.

APPENDIX B

Annotated R Code Used for Statistical Analysis and Creation of Figures

B.1 Introduction

The following will include the R code used to analyze the data and to create the figures presented in this thesis, as well as informal annotation of the commands. Much of what is presented here was only possible because of NRES 598 - SAR: Practical Statistical Analysis with R taught by Dr. David Zaya. I am very grateful for the time and effort Dr. Zaya put into creating this class and for his enthusiasm and willingness to help. While this class was a starting point, there seems to be a steep learning curve for R. Much of the R troubleshooting required in the production of this thesis was carried out using trial and error in conjunction with online forums and various R help websites reached through Google searches. I'm sure there are some places in the following sections where I have done more than is needed to achieve the required result. That was all part of the learning process though! While the small amount of R code presented here doesn't even scratch the surface of what R can do, my hope is that this might in some way aid future members of the Tranel Lab or any others who stumble upon this and may need to perform similar statistics.

B.2 Basics

Some of these figures and data analyses were originally made on a Mac and some on a PC, there may be some differences in the code used for running R on different platforms. The code in following sections was written and run in either R 3.0.3 on my PC or R 3.0.1 on my Mac (R Core Team 2014). While the data sets used throughout this appendix are not available to the reader, the majority of the code in this and the following sections should be able to be used

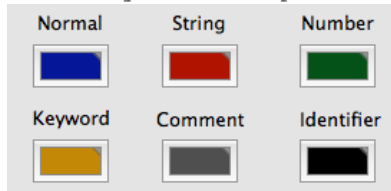
almost verbatim in R if an appropriate similar file is imported. All code is included, but the only output and figures included are those that may be helpful. Comments are provided for almost everything, but may become more condensed as the appendix progresses. Therefore, searching for a particular command of interest throughout the document may be helpful if using only select sections from this guide.

B.2.1 Basic Code

```
# <- comment character - R will not run anything in the line after "#"
# A "#" is placed before all R text output included herein to indicate that the text
is not a command.

# The "<-" symbol is used for naming or assigning
# For example:
mydata<-read.csv("G:\\Thesis\\Multi-state\\Field control data.csv")
# the file read from the location defined in the parentheses above is assigned the
name of "mydata"

# It may be helpful to have all code in one place, rather than typing directly into
the R console
# PC: File --> new script, this opens the R editor
# Mac: File --> new document, opens the text editor in R (The editor on Macs is color-
coded. Colors can be changed under preferences, Syntax. The default colors (Figure
B.1) were used for the code presented in this Appendix, but text color is not
necessary for interpretation.
```



#Figure B.1

```
# Alternatively code can be written in a text editor outside of R and pasted into the
R console

# Shortcut to run a line of code from the open R editor (PC) or document (Mac)
# PC: "ctrl r"
# Mac: "command return"

# Setting the working directory - helpful if working out of the same folder location
for the whole R session
# By doing this, a file can be read from its name only, the entire path doesn't need
to be specified
# PC
setwd(G:\Thesis\Multi-state)
#Mac
setwd("/Volumes/USB20FD/Thesis/Multi-state")

# Check what the working directory is set to
getwd()
```

```

# Importing data into R - data should be saved as .csv or .txt (This can be done from
an excel sheet)
# I am naming all these objects "data" - without naming them R would display the data,
but wouldn't "remember it". By naming it we are able to continue accessing the data.
# PC
mydata<-read.csv("G:\\Thesis\\Multi-state\\Field control data.csv")
mydata<-read.table("G:\\Thesis\\Multi-state\\Field control data.txt", header=TRUE)
# Mac
mydata<-read.csv("/Volumes/USB20FD/Thesis/Multi-state/Field control data.csv")
mydata<-read.table("/Volumes/USB20FD/Thesis/Multi-state/Field control data.txt",
header=TRUE)

# If the working directory is already set to this folder though, the path doesn't need
to be specified
mydata<-read.csv("Field control data.csv")
mydata<-read.table("Field control data.txt", header=TRUE)
# When reading in as a txt, must tell R if there are headers (column names)

# An object stored in R can be viewed by typing its name
Mydata
# This displays the entire data set

# To see the first few lines of the data - shows the names of each column and is good
for see the format of the data
head(mydata)
#      State Rate Pcontrol
#1 Illinois 0.000      0.00
#2 Illinois 0.375     46.53
#3 Illinois 0.750     61.48
#4 Illinois 1.500     75.97
#5 Illinois 3.000     91.51
#6  Kansas 0.000      0.00

# To see the structure of the data
str(mydata)
#'data.frame': 25 obs. of 3 variables:
# $ State : Factor w/ 5 levels "Illinois","Kansas",...: 1 1 1 1 2 2 2 2 2 ...
# $ Rate : num 0 0.375 0.75 1.5 3 0 0.375 0.75 1.5 3 ...
# $ Pcontrol: num 0 46.5 61.5 76 91.5 ...

# Indexing:
# To access only a specific row or column
# What is in the 5th row and the 3rd column of the object "mydata"?
mydata[5,3]

# To look at a particular column you can call it as above
# What is in all rows of the 3rd column?
mydata[,3]
#[1] 91.51

# Or you can index based on the heading in the original file using "$"
mydata$Pcontrol # Displays the "Pcontrol" column from the data set "mydata"
# [1] 0.00 46.53 61.48 75.97 91.51 0.00 48.00 69.00 80.00 96.00 0.00 50.00 65.00
#[14] 87.50 89.80 0.00 13.80 59.50 65.00 90.00 0.00 0.00 0.00 0.00 0.00

# If only working with one data set, it is helpful to "attach" the data, this
eliminates the need for "$"
# If working with multiple data sets, it's easier not to attach and type the path each
time (In this case it's convenient to name the data sets something short)
attach(mydata)

```

```

# Now you can call the 3rd column without specifying the data set to be used
Pcontrol

# Plot the data from the 3rd column
plot(Pcontrol)

# Make a boxplot of the data in the 3rd column
boxplot(Pcontrol)

# After making this second graph, the first disappears!
# Use the following to open a new graphics window, otherwise every new graphic will
replace the previous. Graphics will appear in the "active" graphics window only
windows() # PC
quartz() # Mac

# Installing packages - After a program is installed once, there is no need to re-
install it in subsequent R sessions.
install.packages("RColorBrewer")

# Loading packages - In each R session, a package will have to be loaded prior to use
with the library command
library(RColorBrewer)

# Getting help - You can find documentation on many things by typing a "?" followed by
the function of interest:
?attach
?read.table
?plot
?library

```

B.3 Statistical Analysis and Figures Presented in Chapter 2

The following will include code used in analyzing data from the Illinois statewide survey study and the code used to make the figures presented in Chapter 2. Figure 2.3 was not made in R and therefore is excluded.

B.3.1 Code For Determining a Threshold Copy Number

```

### Importing the copy number data for the sensitive plants and naming it "Sens"
### Importing on a PC
Sens<-read.table("G:\\Thesis\\Statewide Survey\\sensitives.txt", header=TRUE)
### Importing on a Mac
Sens<-read.table("/Volumes/USB20FD/Thesis/Statewide Survey/sensitives.txt",
header=TRUE)

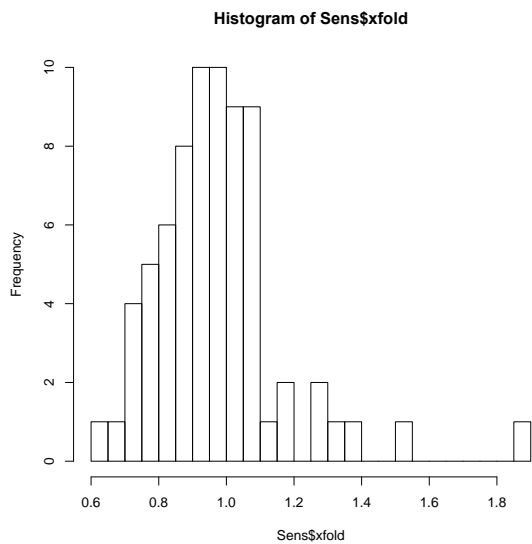
### Showing the first few lines of my data set to see what the columns are called
head(Sens)
#      Sample Index Round xfold
#1 sensitive      1     1  1.02
#2 sensitive      2     1  1.00
#3 sensitive      3     1  0.85
#4 sensitive      4     1  0.60

```



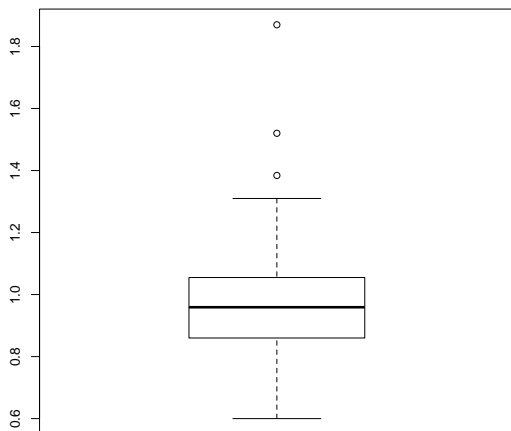
```
#5 sensitive      5      3 1.31
#6 sensitive      6      3 0.89
```

```
### Making a histogram of the copy numbers to see if the data is normally distributed
### breaks refers to "bins" on the x axis
hist(Sens$xfold, breaks=20)
```



#Figure B.2

```
### Make a boxplot of the copy number data
boxplot(Sens$xfold)
```



#Figure B.3

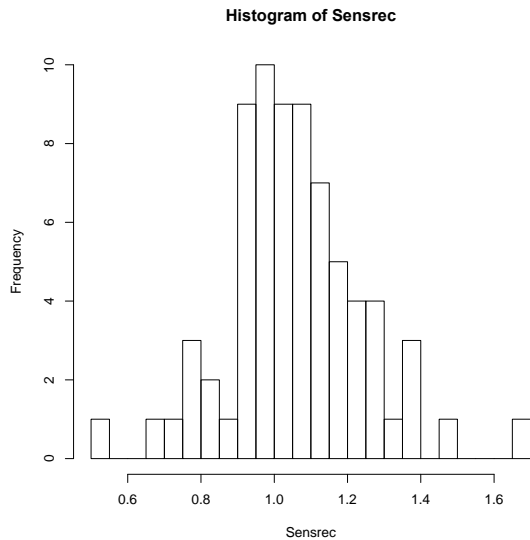
```
### Testing if the data is normal
shapiro.test(Sens$xfold)
# Shapiro-Wilk normality test
#
#data: Sens$xfold
#W = 0.8872, p-value = 9.78e-06
```

```
### Reciprocal transformation on the data and naming the new transformed data set
```

```

"Sensrec"
Sensrec<-1/(Sens$xfold)
hist(Sensrec, breaks=20)

```



#Figure B.4

```
shapiro.test(Sensrec)
```

```

# Shapiro-Wilk normality test
#
#data: Sensrec
#W = 0.9796, p-value = 0.2931

```

```
### Now the data is normal
```

```

mean(Sensrec)
#[1] 1.058181
sd(Sensrec)
#[1] 0.1881803

```

```

### At what copy number are 90%, 95%, and 97% of the data below
### Finding the nth percentile:

```

```

quantile(Sensrec, c(.90,.95,.97))
# 90% 95% 97%
#1.267427 1.379294 1.397590

```

```

1/quantile(Sensrec, c(.90,.95,.97))
# 90% 95% 97%
#0.7890000 0.7250086 0.7155175

```

```

### The above has computed the points at which 90, 95, 97% of the data are above
### Reciprocally transforming flips the data set

```

```

quantile(Sensrec, c(.1,.05,.03))
# 10% 5% 3%
#0.8459083 0.7698680 0.7278494

```

```

1/quantile(Sensrec, c(.1,.05,.03))
# 10% 5% 3%
#1.182161 1.298924 1.373911

```

```

### A conservative cutoff point of 1.4 was chosen. Anything with a relative copy
number above was considered to have elevated EPSPS copy number

```



```

### Saving the point coordinates as a csv, otherwise these coordinates will be lost
and I would have to re-click in all the counties again if returning in a different R
session. Files will be saved in the current working directory.
setwd("/Volumes/USB20FD/Thesis/Statewide Survey") # setting the working directory first
write.csv(res, file="respoints.csv", col.names=TRUE)

### Same as above for the other population types
### Resistant counties w/ gene amplification
resga<-locator(18)
points(x=resga$x, y=resga$y, pch=21, col="black", bg="black", cex=1.4)

### I didn't do a very good job clicking and one of my points was a bit off, you can
index the data set and replace just a specific coordinate without having to re-click
each county
# Here the 10th set of coordinates is being changed slightly
resga$y[10]<-37.25
resga$x[10]<--89.15

### Saving the dataset
write.csv(resga, file="resgapoints.csv", col.names=TRUE)

### Sensitive counties
sen<-locator()
### If you don't specify a number, you can click as many times as you want and then
press "esc" to stop

### Fixing a few points
sen$x[35]<--90.72
sen$y[35]<-39.32

### Checking them - only plotting a point for coordinate set 35
points(x=sen$x[35], y=sen$y[35])

### Plotting all sensitive county points
points(x=sen$x, y=sen$y, pch=4, col="black", bg="gray100", cex=1.4)

### Saving
write.csv(sen, file="senpoints.csv", col.names=TRUE)

#### Later, in a different R session ####
# Because the locations determined using the locator function I the previous session
were saved as csv files, they can be accessed again

### Loading the maps program
library(maps)

### Setting the working directory
setwd("/Volumes/USB20FD/Thesis/Statewide Survey")

### Importing files created during the last session - no county clicking required!
res<-read.csv("respoints.csv")
resga<-read.csv("resgapoints.csv")
sen<-read.csv("senpoints.csv")

### this is what these data sets look like - just like the output from earlier
### typing the name of an object displays it
res
# X x y
#1 1 -90.60242 40.16232
#2 2 -89.90328 40.26820
#3 3 -90.16374 39.71762
#4 4 -89.43718 38.89175

```

```

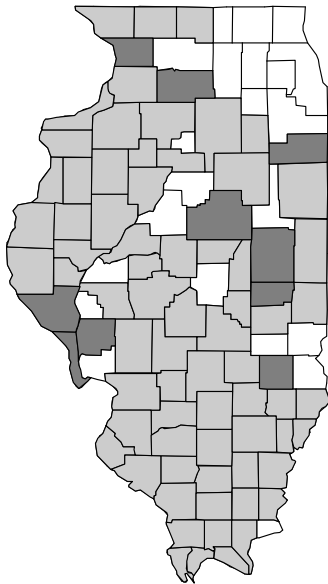
### Coloring specific counties - first list them
IL<-map("county", regions="illinois", plot=FALSE)
IL$names
### This prints a list that looks like:
#[1] "illinois,adams"      "illinois,alexander" "illinois,bond"
#[4] "illinois,boone"      "illinois,brown"     "illinois,bureau"

### Based on this list specific counties can be colored
### First make the starting blank map
IL<-map("county", regions="illinois", plot=TRUE)

### Soybean fields, numbers come from the list created above
map("county",
IL$names[c(1,2,3,5,6,11,13,14,15,20,23,24,25,26,28,29,30,32,33,34,36,37,38,39,41,43,44
,48,50,51,53,54,55,59,60,61,62,63,64,65,66,67,68,69,70,72,73,74,76,77,79,80,81,82,83,8
4,85,87,88,89,91,92,93,94,95,96,97,98,100,101)], fill=TRUE, add=TRUE, col="gray80",
plot=TRUE)

###Corn fields
map("county", IL$names[c(7,8,10,21,31,40,46,52,57,75)], fill=TRUE, add=TRUE,
col="gray50", plot=TRUE)

```



#Figure B.6

```

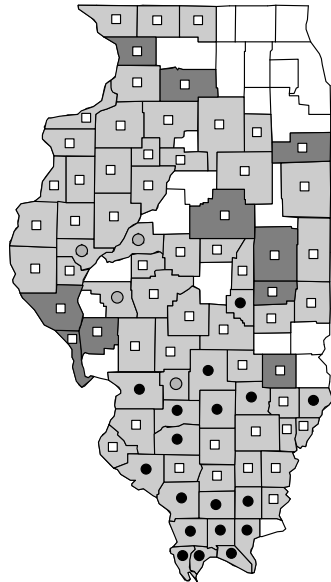
### For more information on the commands used above:
?points
?cex #Character expansion
?pch #Plotting character

### Now adding the points on top (from the csv files created earlier)
points(x=res$x[res$X<3], y=res$y[res$X<3], pch=24,.col="black", bg="black", cex=1.0)
points(x=resga$x, y=resga$y, pch=21, col="black", bg="black", cex=1.2)
points(x=sen$x, y=sen$y,pch=22, col="black",bg="gray100",cex=1.2)

### I needed to change a few counties to a different color - this is easy to do with
indexing
### Morgan county
points(x=res$x[res$X==3], y=res$y[res$X==3],pch=21, col="black",bg="black",cex=1.2)
# here a point is being plotted based on coordinates from the "x" column of the "res"
object, but only when the "X" column of the "res" object is equal to 3. (In other
words, the third set of coordinates only is being plotted)

```

```
### Bond county
points(x=res$x[res$X==4], y=res$y[res$X==4],pch=21, col="black",bg="black",cex=1.2)
```



#Figure B.7

I chose to make the legend for figure 2.1 in powerpoint instead of in R, all other modifications seen in fig. 2.1 were also made in powerpoint

B.3.3 Proportion Bar Graph Code

```
### Importing the data set
### PC
Prop4<-read.table("G:\\Statewide Survey\\Thesis\\Proportion1.4.txt",header=TRUE)
### Mac
Prop4<-read.table("/Volumes/USB20FD/Thesis/Statewide
Survey/Proportion1.4.txt",header=TRUE)

### Installing and loading ggplot2 (only need to install the first time of use)
install.packages("ggplot2")
library(ggplot2)
```

ggplot2 is used here to produce a nice looking graph, for more information:
?ggplot2

Print the first few lines of the data set

```
head(Prop4)
#  above total      prop  county meancopy
#1   10      12 0.8333333 Alexander    3.5
#2    1       7 0.1428571      Bond    1.4
#3   15      15 1.0000000      Clay    3.4
#4    4       5 0.8000000  Clinton    6.9
#5   10      10 1.0000000  Fayette    5.2
#6    8       8 0.8000000  Jackson    3.4
```

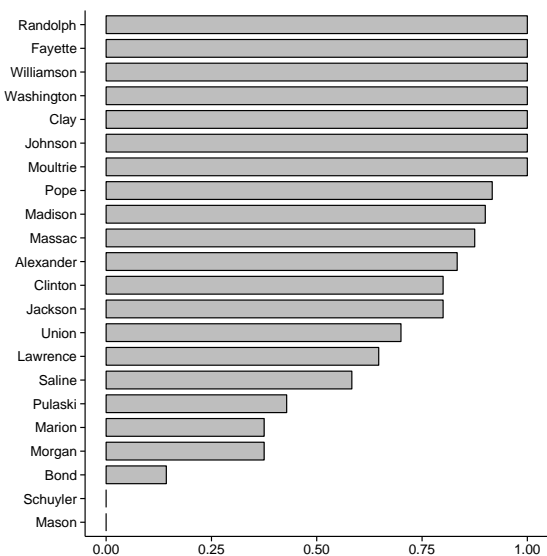
```
### Making the graph using the data "Prop4"
### aes (aesthetics) - county is on the x axis and proportion on the y axis
```



```

### This doesn't look very nice though
### Making the graph again, but changing a few settings
ggplot(Prop4, aes(x=county1,y=prop)) +
  geom_bar(width=0.75,colour="black",fill="gray",
  stat="identity", position="dodge")+ # specifying how the bars should be drawn
  coord_flip()+ # flipping the x and y axis
  theme_bw()+ # changing the theme to black and white
  theme(axis.line=element_line(colour="black"), # making axis black instead of gray
  panel.grid.major=element_blank(), # removing other grid lines
  panel.grid.minor=element_blank(),
  panel.border=element_blank(),
  panel.background=element_blank(),
  axis.title.x=element_blank(), # removing axis titles
  axis.title.y=element_blank(),
  text=element_text(size=15)) # changing text size
# More information ?ggplot2

```

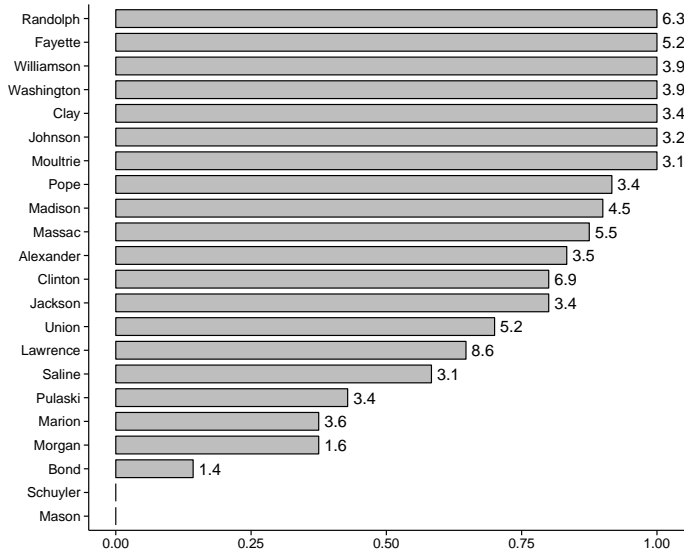


#Figure B.10

```

### Labeling each bar with the mean copy number for that population (using geom_text)
ggplot(Prop4, aes(x=county1,y=prop)) +
  geom_bar(width=0.75,colour="black",fill="gray", stat="identity",
  position="dodge")+
  coord_flip()+
  theme_bw()+
  theme(axis.line=element_line(colour="black"),
  panel.grid.major=element_blank(),
  panel.grid.minor=element_blank(),
  panel.border=element_blank(),
  panel.background=element_blank(),
  axis.title.x=element_blank(),
  axis.title.y=element_blank(),
  text=element_text(size=15))+
  geom_text(aes(label=meancopy),size=5, hjust=-0.25, vjust=0.5) # applies text
using the "meancopy" column, can adjust size and location

```

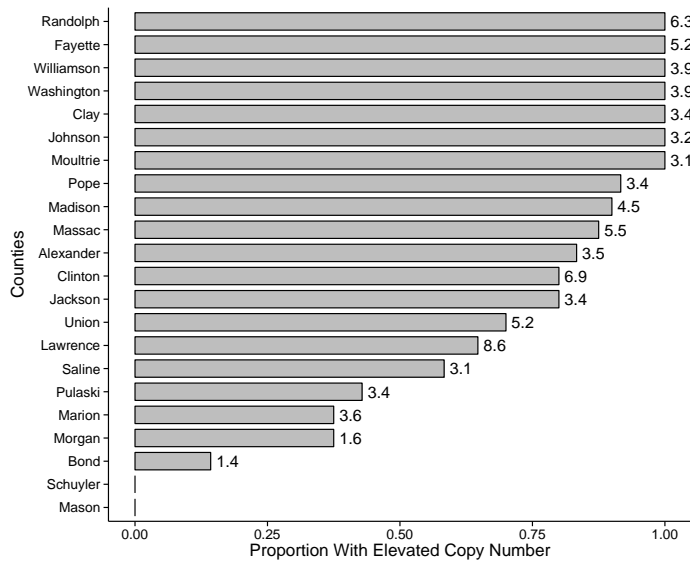



#Figure B.11

```

### Adding axis titles
ggplot(Prop4, aes(x=countyl,y=prop)) +
  geom_bar(width=0.75,colour="black",fill=c("gray"), stat="identity",
position="dodge")+
  coord_flip()+
  theme_bw()+
  xlab("Counties") + ylab("Proportion With Elevated Copy Number") +
  theme(axis.line=element_line(colour="black"),
panel.grid.major=element_blank(),
panel.grid.minor=element_blank(),
panel.border=element_blank(),
panel.background=element_blank(),
text=element_text(size=15))+
  geom_text(aes(label=meancopy),size=5, hjust=-0.25, vjust=0.5)

```



#Figure B.12

B.4 Statistical Analysis and Figures Presented in Chapter 3

The following sub-sections will include the code used to analyze the data and create the figures for the multi-state study presented in Chapter 3. Figure 3.3 was created using Microsoft Excel and Powerpoint; its creation is not discussed here.

B.4.1 Multi-State Data Analysis and Bootstrapping

```
### Importing the data and naming it "FS13"
FS13<-read.table("/Volumes/USB20FD/NRES 598/FS13.txt", header=T)

### Examining the data
summary(FS13)
# Treatment State Foldchange Block
# Min. :0.000 Illinois: 94 Min. : 0.388 G : 34
# 1st Qu.:0.375 Kansas : 80 1st Qu.: 1.071 E : 30
# Median :0.750 Kentucky: 78 Median : 2.013 F : 30
# Mean :1.109 Missouri: 80 Mean : 2.520 L : 25
# 3rd Qu.:1.500 Nebraska:100 3rd Qu.: 3.410 M : 25
# Max. :3.000 Max. :11.507 N : 25
# (Other):263

head(FS13) # Looking at the first few lines of data
# Treatment State Foldchange Block
#1 0.375 Missouri 0.661 A
#2 0.375 Missouri 2.726 A
#3 0.375 Missouri 4.190 A
#4 0.375 Missouri 3.152 A
#5 0.375 Missouri 1.327 B
#6 0.375 Missouri 4.492 B

### Removing the Kentucky Population from the Dataset and calling the new object
"FS13noKY". The KY population did not have gene amplification.
FS13noKY <- FS13[FS13$State != "Kentucky",]
### This can be read as "FS13noKY gets FS13 dataset when the state is NOT Kentucky"

summary(FS13noKY)
# Treatment State Foldchange Block
# Min. :0.000 Illinois: 94 Min. : 0.467 G : 34
# 1st Qu.:0.375 Kansas : 80 1st Qu.: 1.523 E : 30
# Median :0.750 Kentucky: 0 Median : 2.300 F : 30
# Mean :1.112 Missouri: 80 Mean : 2.896 L : 25
# 3rd Qu.:1.500 Nebraska:100 3rd Qu.: 3.871 M : 25
# Max. :3.000 Max. :11.507 N : 25
# (Other):185

### These files have a "block" factor, however I decided not to analyze block because
there are so many levels (4-5 / location = ~25 different blocks) and it is unlikely
that there will be a significant difference between blocks as all plots were located
in the same area of the field.

### Examining the Data

### Getting the RColorBrewer Package, which has pre-made color schemes
install.packages("RColorBrewer")
```

```

library(RColorBrewer)

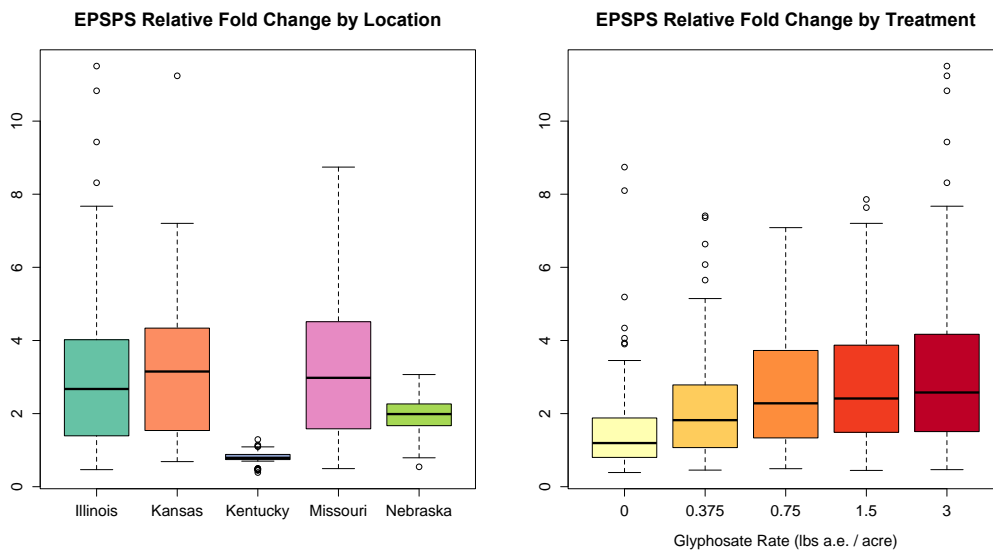
### For more information on RColorBrewer
?RColorBrewer
### To see all the color palettes
display.brewer.all()

### I want to see both of the following boxplots side by side
par(mfcol = c(1, 2))
### This creates a graphing window that allows two graphs side by side, (1 Row, 2
Columns). Then the code for two graphics can be run at once and they appear in the
same graphing window

### Boxplot of foldchange by location - all locations and treatments
boxplot(FS13$Foldchange~FS13$State,
cex.main=1.5, # change character expansion of title
cex.axis=1.25, # change character expansion of axis
col=brewer.pal(5,"Set2"), # color the boxes with 5 colors from the palette called Set2
in RColorBrewer
main="EPSPS Relative Fold Change by Location") # setting title

### Boxplot of foldchange by Treatment - all locations and treatments
boxplot(FS13$Foldchange~FS13$Treatment, cex.main=1.5, cex.lab=1.25,
cex.axis=1.25,col=brewer.pal(5, "YlOrRd"), # character expansion of various elements
of the grph and specifying the color scheme to use
main="EPSPS Relative Fold Change by Treatment", # specifying the main title
xlab="Glyphosate Rate (lbs a.e. / acre)") # specifying an x axis title

```



#Figure B.13

```

### The data is not normally distributed, but it is good to run a linear model and
test the model assumptions anyway

```

```

### Running a 2-way anova and calling the model "m1.full". Don't need to use the $,
the data set to pull from is specified by "data=" in the model
m1.full<-lm(Foldchange~State+Treatment+State:Treatment, data=FS13)

```

```

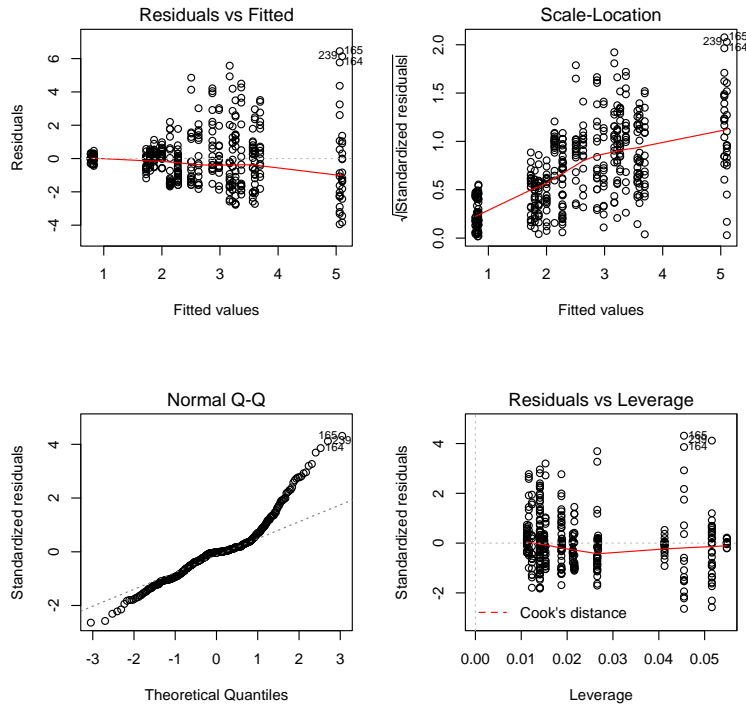
### Looking at the results
summary.aov(m1.full)
#
# State      Df Sum Sq Mean Sq F value    Pr(>F)
# State      4  408.7  102.19  43.646 < 2e-16 ***
# Treatment  1   98.8   98.81  42.204 2.32e-10 ***

```

```
#State:Treatment 4 87.0 21.74 9.285 3.37e-07 ***
#Residuals      422 988.0 2.34
#---
#Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

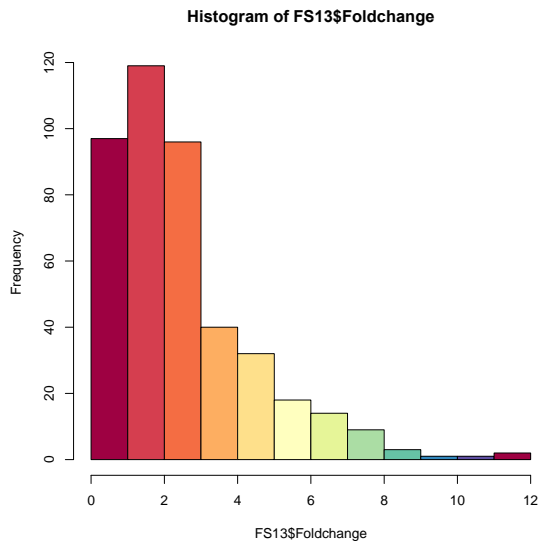
```
### Plotting the model to visually examine normality and homogeneity of variances
par(mfcol = c(2, 2))
### This creates a graphing window that allows four graphs in a 2 x 2 arrangement
```

```
plot(m1.full)
```



#Figure B.14

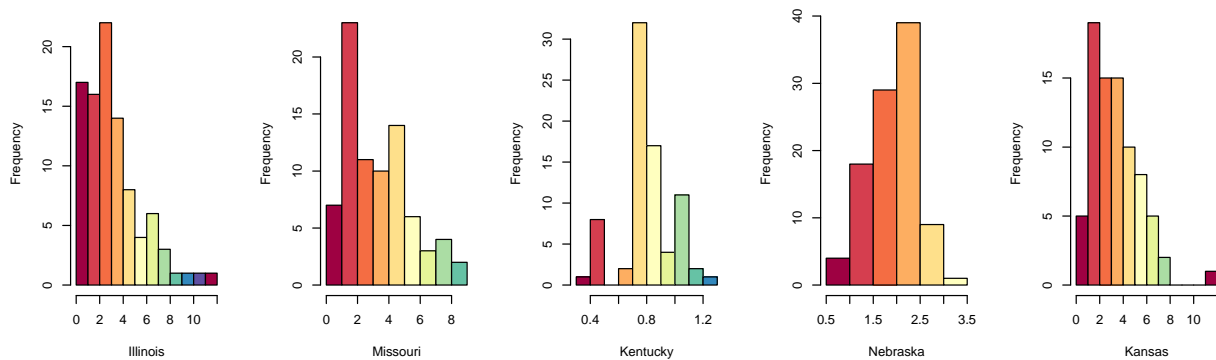
```
### Making a histogram with to look at the distribution of data (color scheme is just
to illustrate RColorBrewer schemes)
hist(FS13$Foldchange, col=brewer.pal(13,"Spectral")) # 13 Colors from the RColorBrewer
"Spectral" color scheme#Figure B.15
```



```

### Histograms by state
par(mfcol = c(1,5)) # Makes a graphing window with 1 row and 5 columns
hist(FS13$Foldchange[FS13$State=="Illinois"], col=brewer.pal(13, "Spectral"),
     xlab="Illinois", main="")
hist(FS13$Foldchange[FS13$State=="Missouri"], col=brewer.pal(13, "Spectral"),
     xlab="Missouri", main="")
hist(FS13$Foldchange[FS13$State=="Kentucky"], col=brewer.pal(13, "Spectral"),
     xlab="Kentucky", main="")
hist(FS13$Foldchange[FS13$State=="Nebraska"], col=brewer.pal(13, "Spectral"),
     xlab="Nebraska", main="")
hist(FS13$Foldchange[FS13$State=="Kansas"], col=brewer.pal(13, "Spectral"),
     xlab="Kansas", main="")

```



#Figure B.16

```

### Running tests to examine normality and homogeneity of variances
shapiro.test(FS13$Foldchange)
### p=2.2e-16
### Data are not normally distributed

```

```

m1.full$residuals
shapiro.test(m1.full$residuals)
### p=3.223e-14
### Reject the null, residuals are not normally distributed

```

```

bartlett.test(FS13$Foldchange, FS13$State)
### p=2.2e-16
### reject null - variances are not equal

```

```

bartlett.test(FS13$Foldchange, FS13$Treatment)
### p=4.919e-06
### reject null - variances are not equal

```

These results suggest that a nonparametric approach is needed to analyze the data
A permutation test doesn't assume normality, but does assume similarity of variances. This makes fewer assumptions than a linear model though

```

### Getting lmPerm package
install.packages("lmPerm")
library(lmPerm)

```

```

### Hypotheses
### H0= Neither state, treatment nor the interaction between state and treatment has
an effect on fold change
### HA= State, treatment and the interaction between state and treatment have an

```

effect on fold change

```
### Permutation test
m1perm<-aovp(Foldchange ~ State + Treatment + State:Treatment, data=FS13)
summary(m1perm)
#Component 1 :
#
#      Df R Sum Sq R Mean Sq Iter  Pr(Prob)
#State      4  410.89  102.723 5000 < 2.2e-16 ***
#Treatment  1   93.85   93.853 5000 < 2.2e-16 ***
#State:Treatment  4   86.96   21.739 5000 < 2.2e-16 ***
#Residuals 422  988.00    2.341
#---
#Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

### Reject null - all three have an effect on fold change

### What happens if KY (the clear outlier, with no gene amplification) is removed?
m2perm<-aovp(Foldchange ~ State + Treatment + State:Treatment, data=FS13noKY)
summary(m2perm)
#Component 1 :
#
#      Df R Sum Sq R Mean Sq Iter  Pr(Prob)
#State      3  131.84   43.947 5000 < 2.2e-16 ***
#Treatment  1  122.59  122.593 5000 < 2.2e-16 ***
#State:Treatment  3   63.89   21.297 5000 < 2.2e-16 ***
#Residuals 346  985.55    2.848
#---
#Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

### Bootstrapping ###

### Bootstrapping the correlation coefficient to look at foldchange by treatment
### All states combined except Kentucky, at all treatment levels

### Defining a function that calculates the correlation coefficient
findr<-function(f){
  cor(f, FS13noKY$Treatment)
}

### Determine how many observations there are in the data being used
length(FS13noKY$Foldchange)
#354

faker<-numeric() #defining a vector called "faker", will be filled later
fakedata<-numeric() #defining a vector called "fakedata", will be filled later

### A "For" loop is used here for resampling many times
### At least 99999 resamples should be used, but this takes a long time, so I used 999
just to check to see if the code works

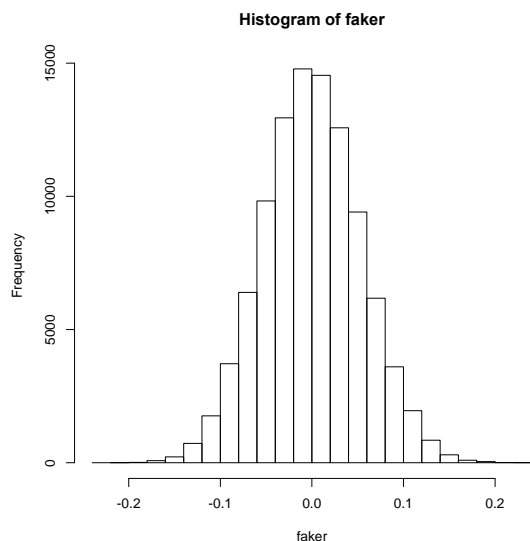
for(i in 1:999){
  fakedata<-sample(FS13noKY$Foldchange, 354, replace=TRUE)
  faker[i] <- findr(fakedata)
}

for(i in 1:99999){
  fakedata<-sample(FS13noKY$Foldchange, 354, replace=TRUE)
  faker[i] <- findr(fakedata)
}

### In each loop, R will resample 354 times with replacement from the specified data
set to create a new resampled fake data set called "fakedata". Then R will determine
the correlation coefficient for the fake data set. These fake r's from each fake data
```

set will be stored in the "faker" vector. After completion of all loops, the "faker" vector will contain 999 r values, or 99999 r values (depending on how many times the loop is run)

```
### Visualizing the distribution of fake r's with a histogram
hist(faker)
```



#Figure B.17

```
### Finding r value in the original (real) data set. This is done two ways below -
first with the cor function and then with the function I defined earlier.
cor(FS13noKY$Foldchange, FS13noKY$Treatment)
#[1] 0.3040766
findr(FS13noKY$Foldchange)
#[1] 0.3040766
```

```
### How many times do we see an r value of 0.3040766 by chance alone? (How many times
is an r value of this or larger observed in the distribution of fake r values? You can
predict that the number will be pretty small given the histogram above since there
aren't many values past this point on the graph)
sum(faker>0.3040766)
#[1] 0
```

```
### To calculate the P-value (this is why using 1 less than a round number makes
things easy)
(0+1)/(99999+1)
#[1] 1e-05
### According to this - gene amplification is associated with glyphosate resistance in
waterhemp
```

```
### Bootstrapping correlation coefficient to look at foldchange by treatment
### Using all states combined except Kentucky, at treatments > 0
### From 0 to 0.375 is really more of a categorical factor - no glyphosate compared to
glyphosate treatment. By taking out the 0 rate I can learn more about the
relationship between EPSPS foldchange and resistance
### If the p value is no longer significant, this suggests that the relationship is
qualitative (for example: populations in which gene amplification is associated with
resistance but applying glyphosate at a higher rate does not select for higher copy
number plants. (A plant with 3 copies will have the same fitness as a plant with 10
copies under 4x glyphosate application)
```

```
### If the p value is still significant, this suggests that the relationship between
EPSPS fold change and glyphosate resistance is quantitative in nature (for example:
populations in which gene amplification is associated with resistance and higher rates
select for plants with increasingly higher EPSPS copy number. (A plant with 3 copies
will have lower fitness than a plant with 10 copies under 4x glyphosate application)
```

```
### Defining new vectors that don't include the 0 treatment
```

```
foldc<-c(FS13noKY$Foldchange[FS13noKY$Treatment>0])
```

```
treat<-c(FS13noKY$Treatment[FS13noKY$Treatment>0])
```

```
state<-c(FS13noKY$State[FS13noKY$Treatment>0])
```

```
### The difference in the data sets can be seen in the boxplots
```

```
par(mfcol = c(1, 2))
```

```
### Boxplot showing foldchange by treatment for all states but KY, and all treatments
```

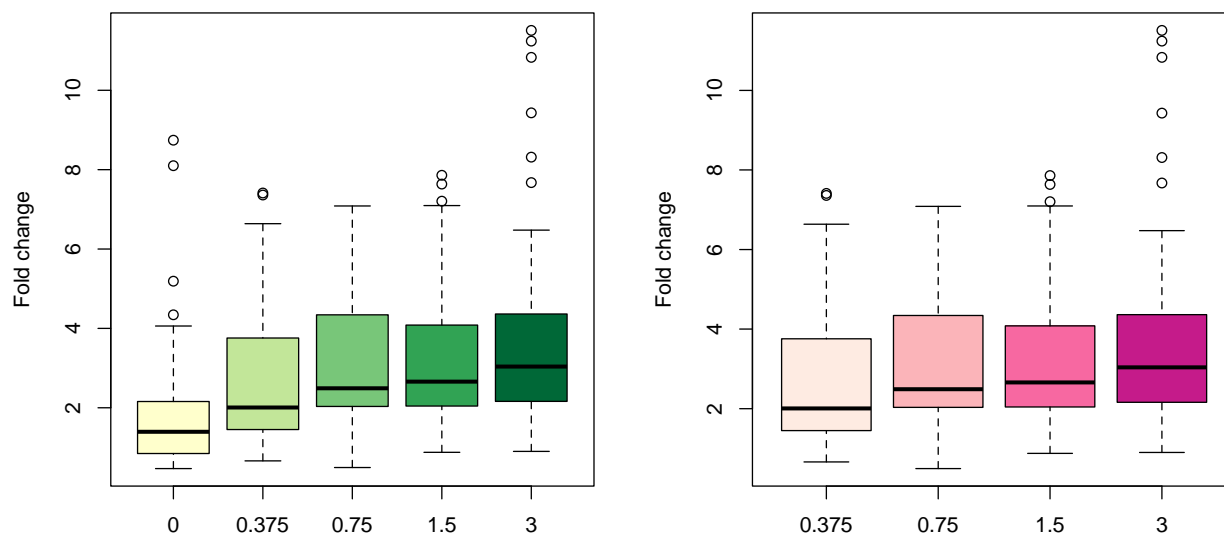
```
boxplot(FS13noKY$Foldchange~FS13noKY$Treatment,
col=brewer.pal(5,"YlGn"), #color boxes with 5 colors from the YlGn palette in
RColorBrewer
```

```
ylab="Fold change") #specify y axis label
```

```
### Boxplot showing foldchange by treatment for all states but KY, and all treatments
except 0
```

```
boxplot(foldc~treat,
col=brewer.pal(5,"RdPu"), #color boxes with 5 colors drawn from the RdPu palette in
RColorBrewer
```

```
ylab="Fold change") #specify y axis label
```



#Figure B.18

```
### Defining function again to find r using treatment values without 0x as the x
```

```
findrnew<-function(f){
```

```
  cor(f, treat)
```

```
}
```

```
### Determine the length of the new data set
```

```
length(foldc)
```

```
#280
```

```
faker1<-numeric() # defining new vectors for this round of bootstrapping
```

```
fakedatal<-numeric()
```

```
for(i in 1:999){
```

```
  fakedatal<-sample(foldc, 280, replace=TRUE)
```

```
  faker1[i] <- findrnew(fakedatal)
```

```
}
```

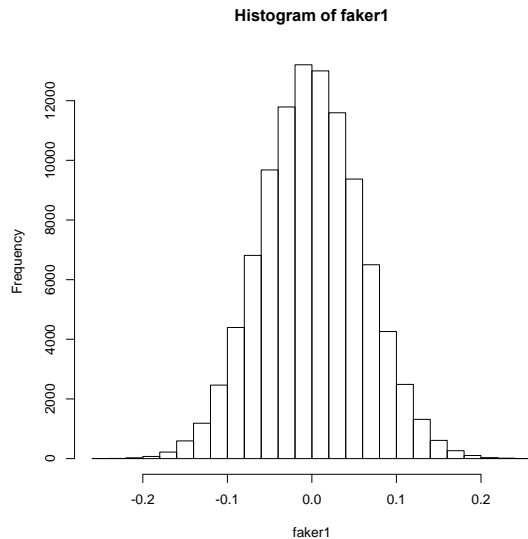


```

for(i in 1:99999){
  fakedata1<-sample(foldc, 280, replace=TRUE)
  faker1[i] <- findrnew(fakedata1)
}

### Looking at the distribution of fake r1 values
hist(faker1)

```



#Figure B.19

```

### Finding the correlation coefficient for the original data set (w/o) the 0x
glyphosate treatment
findrnew(foldc)
cor(foldc, treat)
#[1] 0.1951253

```

```

### How many times do we see an r1 of 0.1951253 by chance alone? (How many times is an
r1 value of this or larger observed in the distribution of fake r1 values? Based on
the histogram above, it looks there are a few fake r1 values above 0.19, but not very
many)
sum(faker1>0.1951253)
#[1] 69

```

```

### Calculating the p-value
(69+1)/(99999+1)
#[1] 7e-04

```

According to this, the relationship between glyphosate rate and relative EPSPS copy number is quantitative in nature. Copy number increases with increasing rate.
 ### This bootstrapping can also be performed individually for each location to look at the relationship at specific locations (not shown)

B.4.2 Map of Study Sites Code

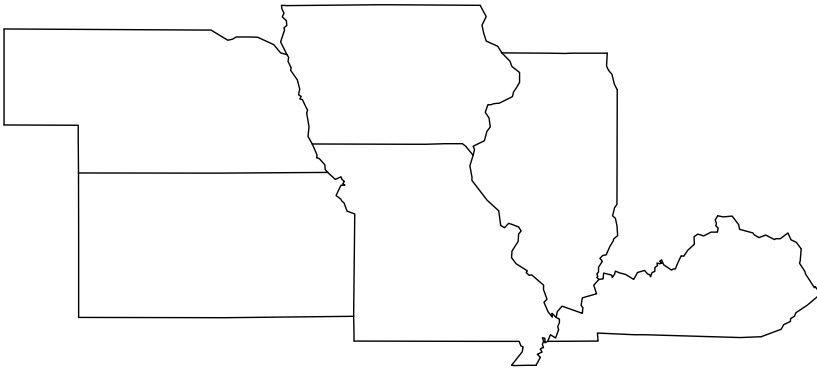
```

### Installing and Loading the maps package
install.packages("maps") #only need to do this once
library(maps)

### Making a map of the states where dose responses were performed and calling it
"multi"
multi<-map('state', regions=c("illinois", "kansas","kentucky","missouri","nebraska",

```

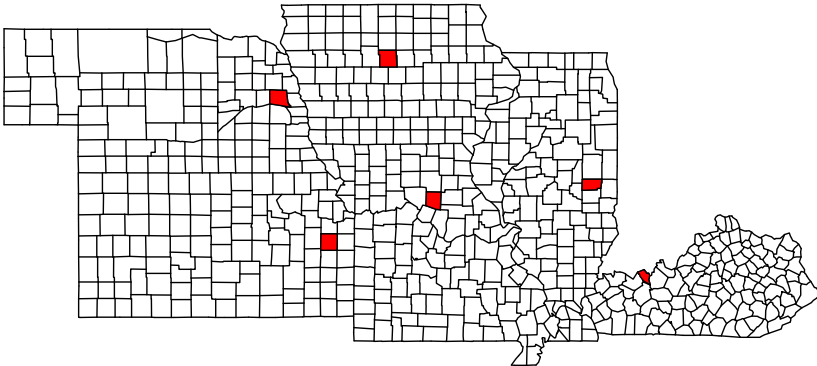
```
"iowa"))
```



#Figure B.20

```
### Print the names associated with this map so I can refer to them later
multi$names
head(multi$names) # viewing just the first few in this list
#[1] "illinois,adams"      "illinois,alexander" "illinois,bond"
#[4] "illinois,boone"     "illinois,brown"    "illinois,bureau"

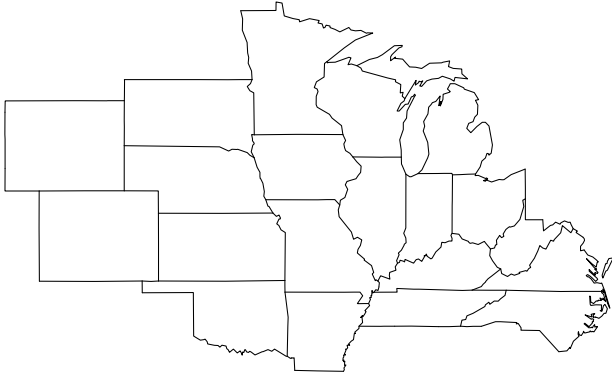
### Making a map of the counties and coloring the ones were studies were performed
multi<-map("county", regions=c("illinois", "kansas","kentucky","missouri","nebraska",
"iowa"))
map("county", multi$names[c(21,142,231,352,514,568)], #these numbers refer to those in
the list made using "multi$names"
fill=TRUE, add=TRUE, col="red", plot=TRUE)
```



#Figure B.21

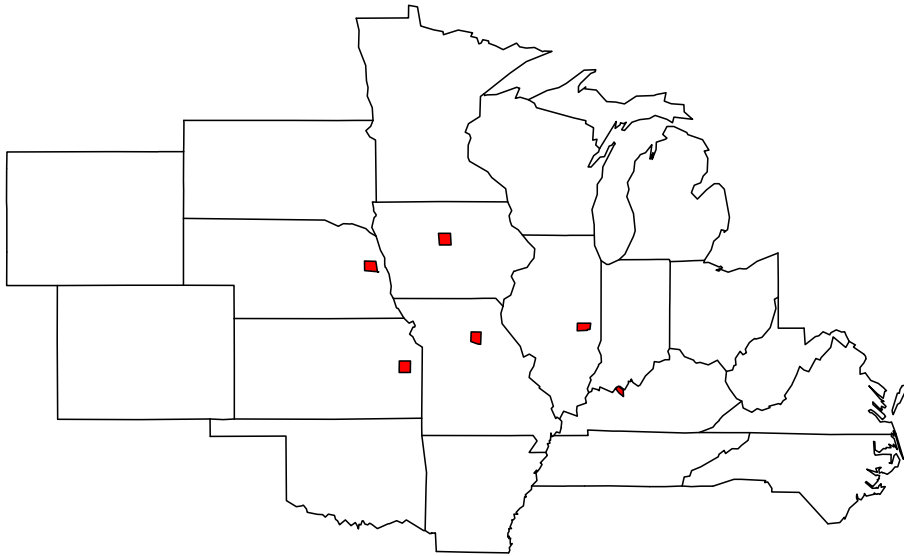
```
### I want some frame of reference (mapping surrounding states) and I don't want all
the counties drawn on the map
### Mapping more counties

map(database="state",
regions=c("missouri","illinois","kansas","illinois","iowa","kentucky","nebraska","arka
nsas","indiana","south
dakota","colorado","wyoming","tennessee","oklahoma","minnesota","wisconsin","ohio","mi
chigan","west virginia","virginia","north carolina"))
```



#Figure B.22

```
### Coloring specific counties
map(database="state",
regions=c("missouri","illinois","kansas","illinois","iowa","kentucky","nebraska","arkansas",
"indiana","south
dakota","colorado","wyoming","tennessee","oklahoma","minnesota","wisconsin","ohio","michigan",
"west virginia","virginia","north carolina"))
map(database="county",
regions=c("nebraska,dodge","illinois,douglas","kansas,franklin","kentucky,hancock","missouri,randolph",
"iowa,hamilton"), add=TRUE, fill=TRUE, col="red")
```



#Figure B.23

```
### Adding the names of each state to the map
### To be sure that the text doesn't overlap the highlighted counties, use the locator
tool - this will allow you to click any where in the graphing window and R will record
the coordinates. Then text can be placed at those coordinates.
### Here the coordinates are saved in an object called "locs"
locs<-locator(6) #Must have the map open in the graphing window for this, curser will
become a "+" in the graphing window. This will allow 6 clicks.

### For unlimited clicks:
locator() # Click as many times as needed, then press esc

locs
#x
```

```

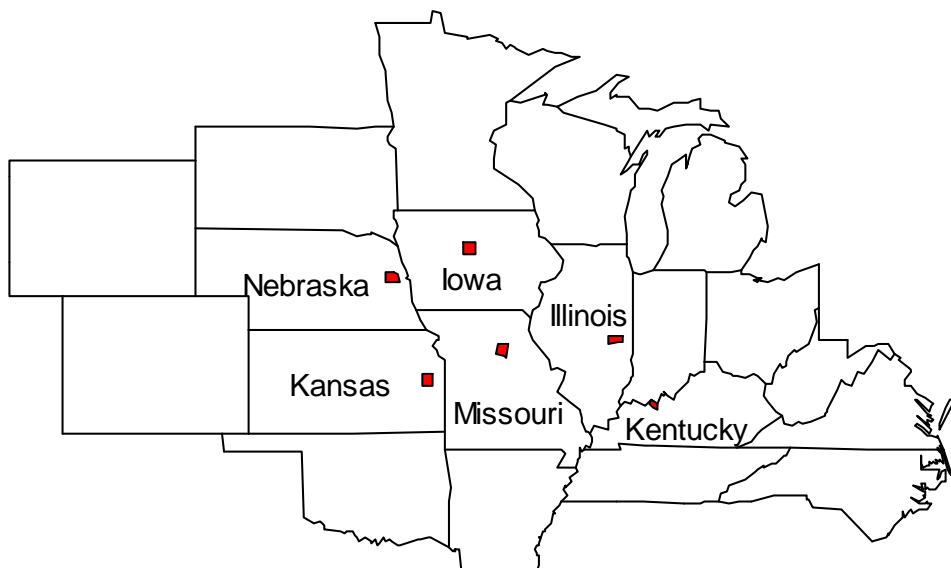
#[1] -89.41010 -93.76665 -98.83935 -85.47130 -92.33436 -99.85388
#
#$y
#[1] 40.71604 41.61870 38.54963 37.19563 37.73723 41.57357

### Saving these coordinates to a file so I can access them in future R sessions
write.csv(locs,file="locs.csv",col.names=TRUE)

### Reading the file back in later
locs<-read.csv("/Volumes/USB20FD/Thesis/Multi-state/locs.csv")

### Adding text to the map
### Labels will be applied to the map in the order that the states were clicked on. I
clicked on the states alphabetically
text(locs$x, locs$y,
labels=c("Illinois","Iowa","Kansas","Kentucky","Missouri","Nebraska"))

```



#Figure B.24

```

### This map was further edited in powerpoint to create the figure seen in Chapter 3

```

B.4.3 Percent Control Ratings Plot Code

```

### Importing the data and calling it "field"
field<-read.table("/Volumes/USB20FD/Thesis/Multi-state/Field control data.txt",
header=T)

head(field)
#   State Rate Pcontrol
#1 Illinois 0.000    0.00
#2 Illinois 0.375   46.53
#3 Illinois 0.750   61.48
#4 Illinois 1.500   75.97
#5 Illinois 3.000   91.51
#6  Kansas 0.000    0.00

### The rates are not metric - making a column within the "field" data set that
contains all rates in metric
field$RateM<-field$Rate*1120

```

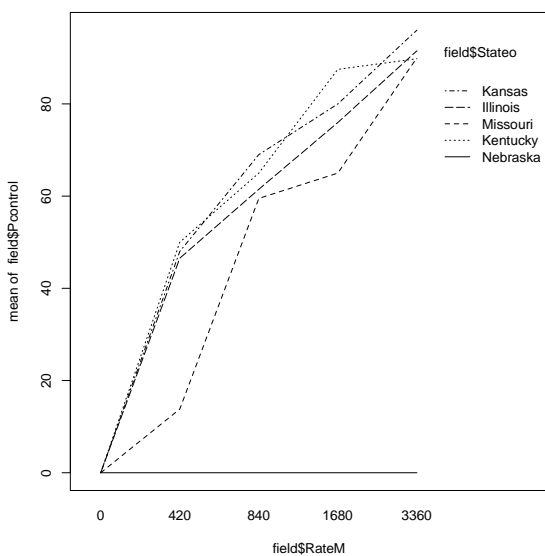
```

### Now the data set has another column called RateM
head(field) #viewing the first few rows
#   State Rate Pcontrol RateM
#1 Illinois 0.000    0.00    0
#2 Illinois 0.375   46.53   420
#3 Illinois 0.750   61.48   840
#4 Illinois 1.500   75.97  1680
#5 Illinois 3.000   91.51  3360
#6  Kansas  0.000    0.00    0

### Making the graph
par(mar=c(5,5,1,1)) # This changes the margins, par(mar=c(bottom, left, top, right))

### I use the interaction plot here just because it was easier for me to plot all 5
states. This function plots the mean, but I only have 1 value for each state and rate.
interaction.plot(field$RateM, field$State, field$Pcontrol)

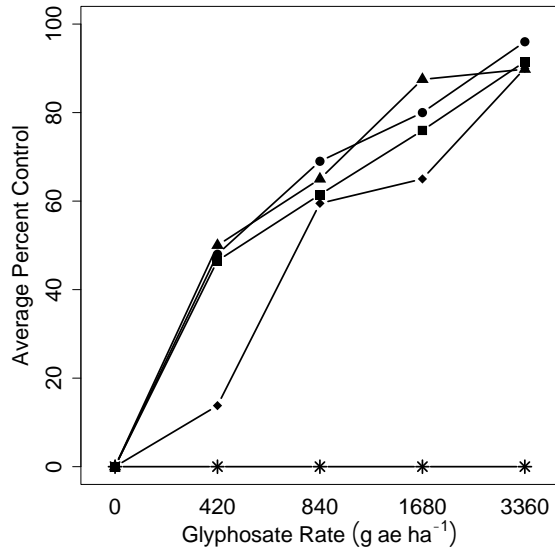
```



```

#Figure B.25
### Changing some of the graphing parameters
interaction.plot(field$RateM, field$State, field$Pcontrol, # plots the different
states percent control values with increasing rate
  cex.axis=1.6, # character expansion of the axis
  type="b", # plots both lines and points (default is line only)
  pch=c(15, 16, 17, 18, 8), # plotting character ?pch for options
  cex.lab=1.6, # character expansion of the labels
  lty=1, lwd=2, # line type and line width
  legend=FALSE, # interaction.plot automatically makes a legend, but I want to make
my own, default is legend=TRUE
  ylab="Average Percent Control",
  cex=1.5, # character expansion
  xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})), # this creates an exponent
  ylim=c(0,100)) # changing the y axis range

```



#Figure B.26

```
### Adding a legend
```

```
### Using locator to obtain coordinates on the graph where clicked
```

```
locator(1) # allows one click on the graph made above and returns coordinates
```

```
#$x
```

```
#[1] 0.7393636
```

```
#
```

```
#$y
```

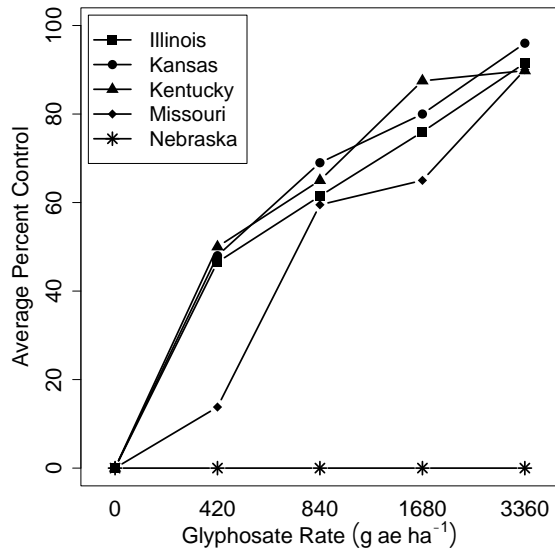
```
#[1] 102.3467
```

```
### Making the legend and placing it where I just clicked
```

```
legend(0.7393636, 102.3467,
```

```
legend=c("Illinois", "Kansas", "Kentucky", "Missouri", "Nebraska"), # text listed in the legend
```

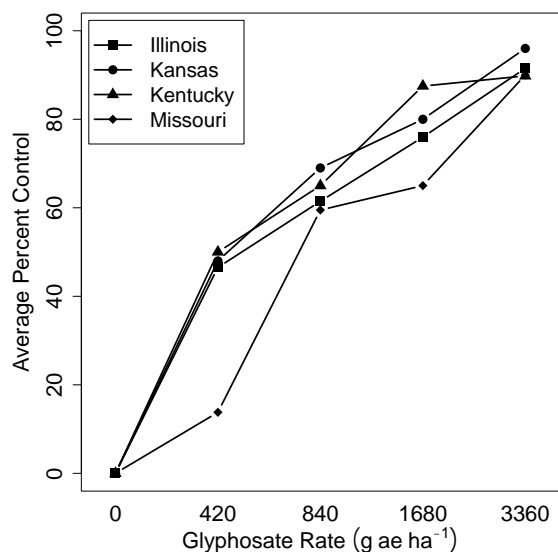
```
cex=1.5, lwd=2, lty=1, pch=c(15, 16, 17, 18, 8)) # these paramaters should be exactly the same as those in the code used to make the graph
```



#Figure B.27

```
### I still haven't heard back from Nebraska about their percent control ratings - I
put 0's in as placeholders originally, but now I want to remove this line since I
don't have the correct values
```

```
### Removing Nebraska - code is almost exactly the same
par(mar=c(5,5,1,1))
interaction.plot(field$RateM[field$State!="Nebraska"],
field$State[field$State!="Nebraska"], field$Pcontrol[field$State!="Nebraska"],
# indexing to plot all but Nebraska, field$RateM[field$State!="Nebraska"] refers to
"object RateM when the corresponding state is anything but Nebraska
  cex.axis=1.6, type="b", pch=c(15, 16, 17, 18), #deleted the pch for Nebraska
  cex.lab=1.6, lty=1, lwd=2, legend=FALSE, ylab="Average Percent Control",
  cex=1.5, xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})), ylim=c(0,100))
legend(0.7393636, 102.3467, legend=c("Illinois", "Kansas", "Kentucky", "Missouri"),
cex=1.5, lwd=2, lty=1, pch=c(15, 16, 17, 18, 8))
```



#Figure B.28

B.4.4 Code For Figure 3.4 –Boxplots, Interaction Plots, and Violin Plots

```
### Importing the data and calling it "FS13"
FS13<-read.table("/Volumes/USB20FD/NRES 598/FS13.txt", header=T)

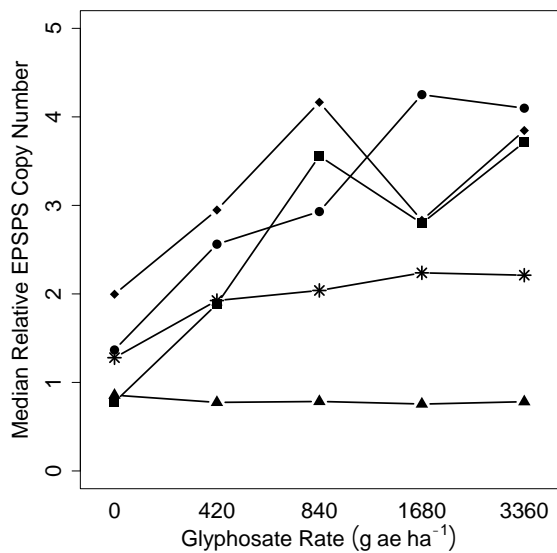
### Examining the first few rows to see format and headings
head(FS13)
# Treatment State Foldchange Block
#1 0.375 Missouri 0.661 A
#2 0.375 Missouri 2.726 A
#3 0.375 Missouri 4.190 A
#4 0.375 Missouri 3.152 A
#5 0.375 Missouri 1.327 B
#6 0.375 Missouri 4.492 B

### The rates aren't metric - adding another column to the data set with metric rates
FS13$TreatmentM<-FS13$Treatment*1120
head(FS13)
# Treatment State Foldchange Block TreatmentM
#1 0.375 Missouri 0.661 A 420
#2 0.375 Missouri 2.726 A 420
#3 0.375 Missouri 4.190 A 420
#4 0.375 Missouri 3.152 A 420
#5 0.375 Missouri 1.327 B 420
#6 0.375 Missouri 4.492 B 420
```

```

### Making an interaction plot to see how the states behave at each rate
par(mar=c(5,5,1,1)) #Changing the margins to allow for more room on the bottom and
left side
interaction.plot(FS13$TreatmentM, FS13$State, FS13$Foldchange,
  fun=median, #plot using the function median (default is mean)
  cex.axis=1.6, #character expansion of the axis
  type="b", #type = both - plot both points and lines
  cex.lab=1.6, #character expansion of the labels
  lty=1, #line type
  lwd=2, #line width
  legend=FALSE, #default is TRUE, doing this so a legend can be added separately
  ylab="Median Relative EPSPS Copy Number", #labeling the y axis
  cex=1.5, #character expansion
  pch=c(15, 16, 17, 18, 8), #plotting character
  xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})), #labeling the x axis using
an exponent
  ylim=c(0,5)) #setting the min and max for the y axis

```



#Figure B.29

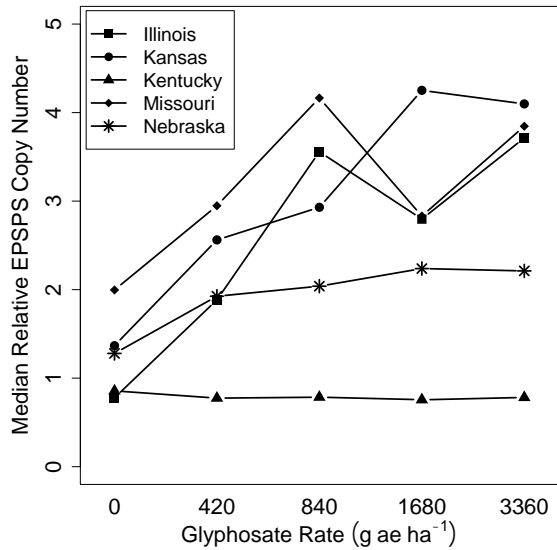
```

### Adding a legend

### Using locator to find the coordinates of a spot clicked in the graphing window
locator(1) # Allows 1 click in the graphing window and returns coordinates
#$x
#[1] 0.7240777
#
#$y
#[1] 5.151097

### Putting the legend in the spot just clicked
legend(0.7240777, 5.151097,
  legend=c("Illinois", "Kansas", "Kentucky", "Missouri", "Nebraska"),
  cex=1.4, lwd=2, lty=1, pch=c(15, 16, 17, 18, 8)) # parameters in the legend should
match those used to make the graph

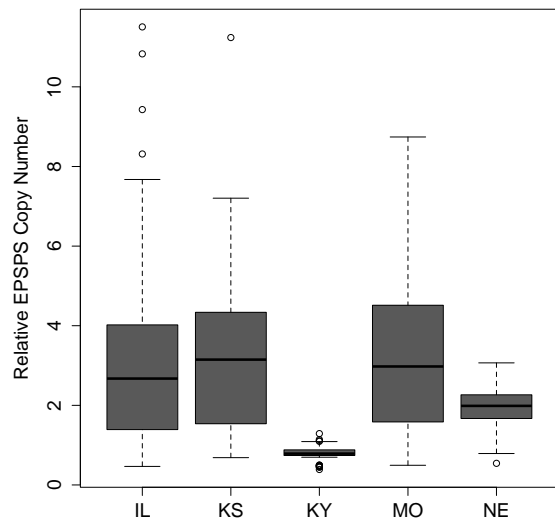
```

#Figure B.30

```
### Making a boxplot - by state
par(mar=c(5,5,1,1)) # setting margins
boxplot(FS13$Foldchange~FS13$State, # making a boxplot showing copy number
distributions of each state
col=c("gray35"), # specifying color
cex.main=1.5, cex.lab=1.4,cex.axis=1.4, # changing the character expansion
xaxt="n", # removing the x axis, by default the full state names are displayed (what
is in the data set), but they dont all fit on the x axis so I'll add my own names.
ylab="Relative EPSPS Copy Number") # labeling the y axis

# Adding x axis labels
axis(1, at=1:5, labels=c("IL", "KS", "KY", "MO", "NE"), cex.axis=1.4)
```



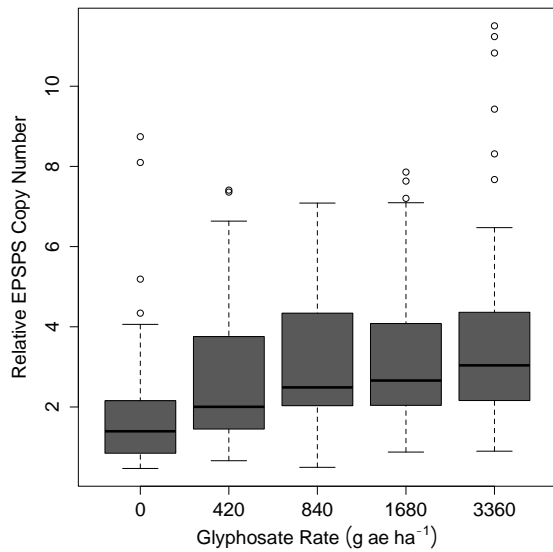
#Figure B.31

```
### Making a boxplot with treatments - very similar to above
par(mar=c(5,5,1,1))
boxplot(FS13$Foldchange[FS13$State!="Kentucky"]~FS13$TreatmentM[FS13$State!="Kentucky"
], # showing the distribution of copy numbers at each rate, while removing all
observations from Kentucky (Kentucky didn't have EPSPS gene amplification)
```

```

col=c("gray35"),
cex.main=1.5, cex.lab=1.4, cex.axis=1.4,
ylab="Relative EPSPS Copy Number",
xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})))

```



#Figure B.32

```

### Making Violin plots - These combine a boxplot with a kernel density plot

```

```

### Installing and loading vioplot
install.packages("vioplot")
library(vioplot)

```

```

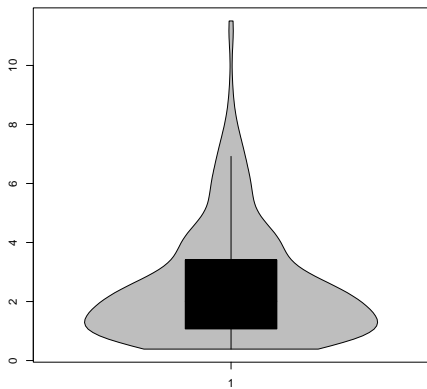
### For more information on vioplot
?vioplot

```

```

### Making a violin plot
vioplot(FS13$Foldchange, col="gray")
### this only plots 1, not a different one for each state - to add all 5 a plot needs
to first be called and then violin plots for each state can be added

```

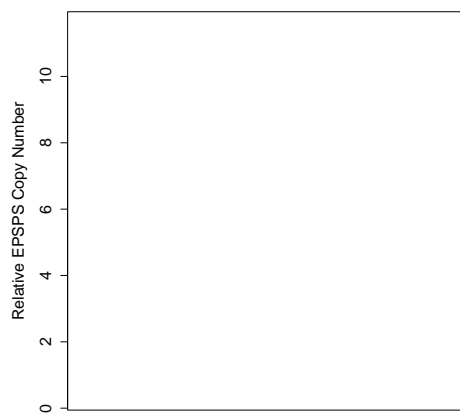


#Figure B.33

```

### Setting up the plotting window
par(mar=c(5,5,1,1)) # setting the plot margins to allow more room on the bottom and
side (order is(bottom, left, top, right))
plot(1,1, # will plot the point (1,1)
xlim=c(0.5,5.5), # I'm going to have 5 states - this range will leave a little bit of
extra space on each side
ylim=range(FS13$Foldchange), # not sure what the y axis range should be, R can do that
for me by telling it that the y axis should be the range of the foldchange values
ylab="Relative EPSPS Copy Number", # labeling the y axis
xlab="", # I don't want an x axis, but R puts one by default - this prevents that
xaxt="n", # This removes the x axis - I'll label it myself with the state
abbreviations
type="n", # There doesn't actually need to be a point at (1,1) - the purpose of this
plot is just to make the plotting window. type="n" means nothing is actually plotted
cex.main=1.5, # character expansion of main title
cex.lab=1.4, # character expansion of labels
cex.axis=1.4) # character expansion of axis

```

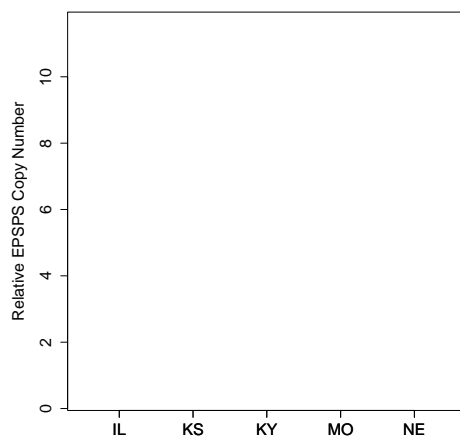


#Figure B.34

```

### Adding my own x axis labels, R would use what is in the data by default which is
the full state name (These don't fit very well.)
axis(1, at=1:5, # place the labels where 1, 2, 3, 4, and 5 would have been
labels=c("IL", "KS", "KY", "MO", "NE"), # specifying my labels
cex.axis=1.4) #character expansion

```



#Figure B.35

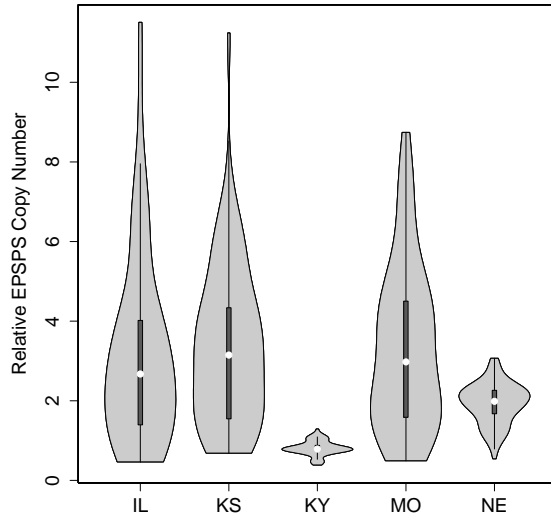
```

### Now adding the violin plots for each state - must add these individually
vioplot(FS13$Foldchange[FS13$State=="Illinois"], # plotting the foldchange
distribution when the state is Illinois
col="gray80", # color of kernel density plot

```

```
add=TRUE, # specify TRUE so that it can be added to the previously made plot window
at=1, # add this plot where 1 would have been on the x axis
rectCol="gray35") # color of the boxplot inside the density plot
```

```
### Same, but with the other states
vioplot(FS13$Foldchange[FS13$State=="Kansas"], col="gray80", add=TRUE, at=2,
rectCol="gray35")
vioplot(FS13$Foldchange[FS13$State=="Kentucky"], col="gray80", add=TRUE, at=3,
rectCol="gray35")
vioplot(FS13$Foldchange[FS13$State=="Missouri"], col="gray80", add=TRUE, at=4,
rectCol="gray35")
vioplot(FS13$Foldchange[FS13$State=="Nebraska"], col="gray80", add=TRUE, at=5,
rectCol="gray35")
```



#Figure B.36

```
### I didn't like the dots in the middle of the boxplots - they don't look like
typical boxplots. I also didn't like how thin they were. With the help of Google
search I found something called "Vioplot2" - an updated code from Daniel Adler (author
of vioplot) that allows lines at the median points instead of dots, and allows
plotting the kernel density plot on a single side (code for vioplot2 available:
https://gist.githubusercontent.com/mbjoseph/5852613/raw/eb6dbf8d79fc69982b6e45a72177d73dc549ecb3/vioplot2)
```

```
### I just slightly altered this code to change the width of my boxplots (my
alterations are marked with comments)
vioplot2 <- function(x, ..., range = 1.5, h = NULL, ylim = NULL, names = NULL,
                    horizontal = FALSE, col = "magenta", border = "black", lty = 1,
                    lwd = 1, rectCol = "black", colMed = "white", pchMed = 19,
                    at, add = FALSE, wex = 1, drawRect = TRUE, side="both")
{
  datas <- list(x, ...)
  n <- length(datas)
  if (missing(at))
    at <- 1:n
  upper <- vector(mode = "numeric", length = n)
  lower <- vector(mode = "numeric", length = n)
  q1 <- vector(mode = "numeric", length = n)
  q2 <- vector(mode = "numeric", length = n)
  q3 <- vector(mode = "numeric", length = n)
  med <- vector(mode = "numeric", length = n)
  base <- vector(mode = "list", length = n)
  height <- vector(mode = "list", length = n)
  baserange <- c(Inf, -Inf)
```

```

args <- list(display = "none")
radj <- ifelse(side == "right", 0, 1)
ladj <- ifelse(side == "left", 0, 1)
if (!is.null(h))
  args <- c(args, h = h)
med.dens <- rep(NA, n)
for (i in 1:n) {
  data <- datas[[i]]
  data.min <- min(data)
  data.max <- max(data)
  q1[i] <- quantile(data, 0.25)
  q2[i] <- quantile(data, 0.5)
  q3[i] <- quantile(data, 0.75)
  med[i] <- median(data)
  iqd <- q3[i] - q1[i]
  upper[i] <- min(q3[i] + range * iqd, data.max)
  lower[i] <- max(q1[i] - range * iqd, data.min)
  est.xlim <- c(min(lower[i], data.min), max(upper[i],
                                             data.max))
  smout <- do.call("sm.density", c(list(data, xlim = est.xlim,
                                       args))
  med.dat <- do.call("sm.density",
                    c(list(data, xlim=est.xlim,
                          eval.points=med[i], display = "none")))
  med.dens[i] <- med.dat$estimate
  hscale <- 0.4/max(smout$estimate) * wex
  base[[i]] <- smout$eval.points
  height[[i]] <- smout$estimate * hscale
  med.dens[i] <- med.dens[i] * hscale
  t <- range(base[[i]])
  baserange[1] <- min(baserange[1], t[1])
  baserange[2] <- max(baserange[2], t[2])
}
if (!add) {
  xlim <- if (n == 1)
    at + c(-0.5, 0.5)
  else range(at) + min(diff(at))/2 * c(-1, 1)
  if (is.null(ylim)) {
    ylim <- baserange
  }
}
if (is.null(names)) {
  label <- 1:n
}
else {
  label <- names
}
boxwidth <- 0.25 * wex # changed 0.05 to 0.25 to increase the width of the boxplots
if (!add)
  plot.new()
if (!horizontal) {
  if (!add) {
    plot.window(xlim = xlim, ylim = ylim)
    axis(2)
    axis(1, at = at, label = label)
  }
  box()
  for (i in 1:n) {
    polygon(x = c(at[i] - radj*height[[i]], rev(at[i] + ladj*height[[i]])),
           y = c(base[[i]], rev(base[[i]])),
           col = col, border = border,
           lty = lty, lwd = lwd)
    if (drawRect) {

```

```

        lines(at[c(i, i)], c(lower[i], upper[i]), lwd = lwd,
              lty = lty)
        rect(at[i] - radj*boxwidth/2,
              q1[i],
              at[i] + ladj*boxwidth/2,
              q3[i], col = rectCol)
        # median line segment
        lines(x = c(at[i] - 0.5*boxwidth, # changed this so that median lines only
                    extend the width of the boxplot
                    at[i],
                    at[i] + 0.5*boxwidth[i]), # changed this so that median lines only extend
                    the width of the boxplot
              y = rep(med[i],3))
    }
}
}
else {
  if (!add) {
    plot.window(xlim = ylim, ylim = xlim)
    axis(1)
    axis(2, at = at, label = label)
  }
  box()
  for (i in 1:n) {
    polygon(c(base[[i]], rev(base[[i]])),
            c(at[i] - radj*height[[i]], rev(at[i] + ladj*height[[i]])),
            col = col, border = border,
            lty = lty, lwd = lwd)
    if (drawRect) {
      lines(c(lower[i], upper[i]), at[c(i, i)], lwd = lwd,
            lty = lty)
      rect(q1[i], at[i] - radj*boxwidth/2, q3[i], at[i] +
            ladj*boxwidth/2, col = rectCol)
      lines(y = c(at[i] - radj*med.dens[i],
                  at[i],
                  at[i] + ladj*med.dens[i]),
            x = rep(med[i],3))
    }
  }
}
invisible(list(upper = upper, lower = lower, median = med,
              q1 = q1, q3 = q3))
}

```

Now I can make the violin plots again using "vioplot2" instead of "vioplot"

Calling a plotting window

```

par(mar=c(5,5,1,1))
plot(1,1,xlim=c(0.5,5.5), ylim=range(FS13$Foldchange), ylab="Relative EPSPS Copy
Number", xlab="", xaxt="n", type="n", cex.main=1.5, cex.lab=1.4, cex.axis=1.4)

```

Adding my own x axis labels

```

axis(1, at=1:5, labels=c("IL", "KS", "KY", "MO", "NE"), cex.axis=1.4)

```

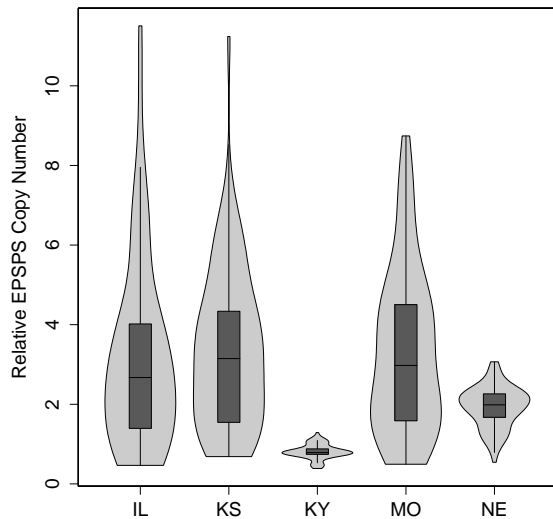
Now adding the violin plots for each state - must add these individually

```

vioplot2(FS13$Foldchange[FS13$State=="Illinois"], col="gray80", add=TRUE, at=1,
rectCol="gray35")
vioplot2(FS13$Foldchange[FS13$State=="Kansas"], col="gray80", add=TRUE, at=2,
rectCol="gray35")
vioplot2(FS13$Foldchange[FS13$State=="Kentucky"], col="gray80", add=TRUE, at=3,
rectCol="gray35")

```

```
vioplot2(FS13$Foldchange[FS13$State=="Missouri"], col="gray80", add=TRUE, at=4,
rectCol="gray35")
vioplot2(FS13$Foldchange[FS13$State=="Nebraska"], col="gray80", add=TRUE, at=5,
rectCol="gray35")
```

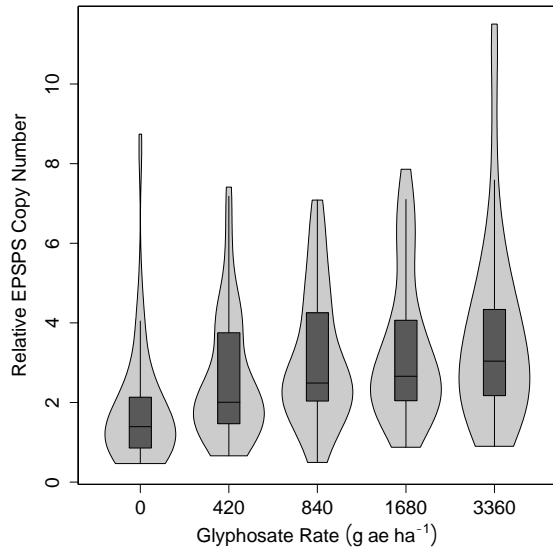


#Figure B.37

```
### Violin plots of copy number by rate
### This is almost exactly the same as previous, but now I've excluded KY

### Making a new data set without KY and converting the treatment rates to metric
FS13noKY <- FS13[FS13$State != "Kentucky",]
FS13noKY$TreatmentM<-FS13noKY$Treatment*1120

par(mar=c(5,5,1,1))
plot(1,1, xlim=c(0.5,5.5), ylim=range(FS13$Foldchange), ylab="Relative EPSPS Copy
Number", xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})),xaxt="n",
type="n",cex.main=1.5, cex.lab=1.4,cex.axis=1.4)
axis(1, at=1:5, labels=c("0", "420", "840", "1680", "3360"), cex.axis=1.4)
vioplot2(FS13noKY$Foldchange[FS13noKY$TreatmentM==0], col="gray80", add=TRUE, at=1,
rectCol="gray35")
vioplot2(FS13noKY$Foldchange[FS13noKY$TreatmentM==420], col="gray80", add=TRUE, at=2,
rectCol="gray35")
vioplot2(FS13noKY$Foldchange[FS13noKY$TreatmentM==840], col="gray80", add=TRUE, at=3,
rectCol="gray35")
vioplot2(FS13noKY$Foldchange[FS13noKY$TreatmentM==1680], col="gray80", add=TRUE, at=4,
rectCol="gray35")
vioplot2(FS13noKY$Foldchange[FS13noKY$TreatmentM==3360], col="gray80", add=TRUE, at=5,
rectCol="gray35")
```



#Figure B.38

B.4.5 Stacked Bar Graph of Pro106Ser Percentages

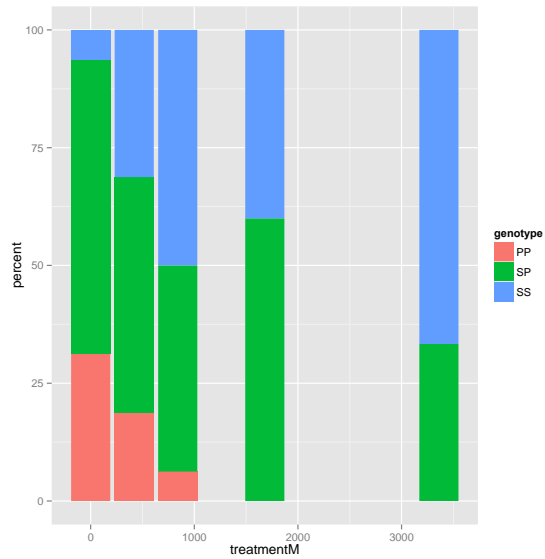
```
### Importing the data and calling it "ProSer"
ProSer<-read.table("/Volumes/USB20FD/NRES 598/ProSerPercentages.txt", header=T)

### Installing and loading ggplot
install.packages("ggplot2") #only need to do this once per R program
library(ggplot2)

### Looking at the first few rows of the data
head(ProSer)
# treatment genotype percent
#1      0x      PP  31.25
#2      0x      SP  62.50
#3      0x      SS   6.25
#4     0.5x     PP  18.75
#5     0.5x     SP  50.00
#6     0.5x     SS  31.25

### The treatment values aren't metric, this adds a column to the data set with the
appropriate rates
ProSer$treatmentM<-c(0,0,0,420,420,420,840,840,840,1680,1680,1680,3360,3360,3360)

### Making the graph
ggplot(data=ProSer, aes(x=treatmentM, y=percent))+ # Specifying what data set to use,
and defining x and y axes
geom_bar(aes(fill=genotype), stat="identity") # using a bar graph geometry
(?geom_bar), and plotting bars filled based on genotype (this is what makes the graph
stacked - there are multiple categories (PP, SP, SS) for each treatment)
```

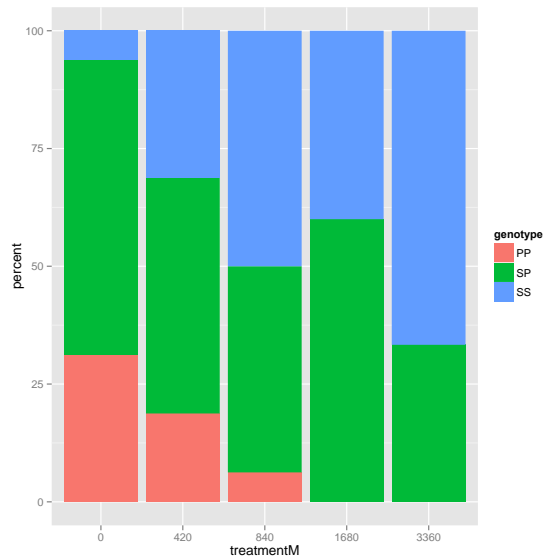
#Figure B.39

I want to remove the spaces between the bars - there is space because the x axis is numeric:

```
class(ProSer$treatmentM)
#[1] "numeric"
```

```
### To make ProSer$treatmentM a factor
ProSer$treatmentM<-as.factor(ProSer$treatmentM)
class(ProSer$treatmentM)
#[1] "factor"
```

```
### Plotting again
ggplot(data=ProSer, aes(x=treatmentM,
y=percent))+geom_bar(aes(fill=genotype), stat="identity")
```



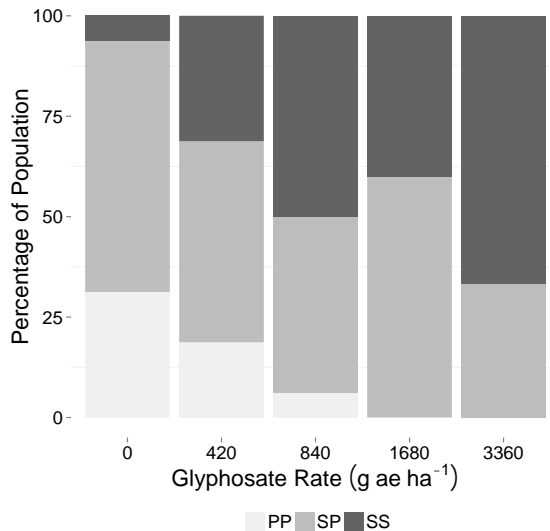
#Figure B.40

```
### Changing some of the parameters
ggplot(data=ProSer, aes(x=treatmentM, y=percent)) +
geom_bar(aes(fill=genotype), stat="identity") +
  xlab(expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1}))) + # Labeling the x axis
  ylab("Percentage of Population")+ # Labeling the y axis
```

```

scale_fill_brewer(palette="Greys") + # Using the gray palette from RColorBrewer
theme(plot.background=element_rect(fill="white"),
panel.background=element_rect(fill="white"), # Making background white
axis.text=element_text(colour="black"), # Making axis text black
text=element_text(size=20), # Changing font size
legend.position="bottom", # Positioning legend under graph
legend.title=element_blank() # Remove legend title

```

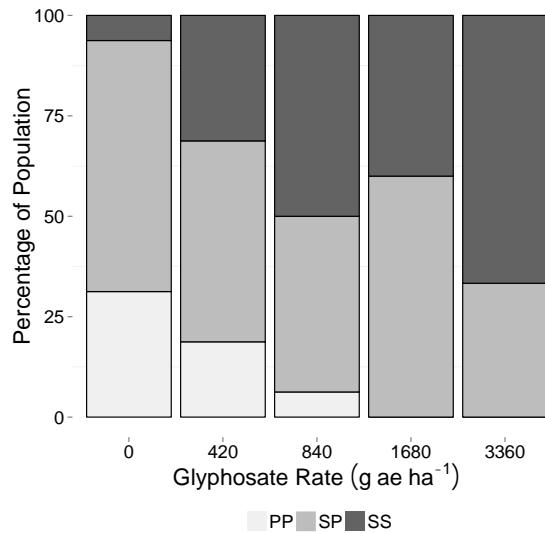


#Figure B.41

```

### I want to outline the bars, but not the legend
### This makes the same graph as above
ggplot(data=ProSer, aes(x=treatmentM, y=percent)) +
geom_bar(aes(fill=genotype), stat="identity") +
xlab(expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1}))) +
ylab("Percentage of Population")+
scale_fill_brewer(palette="Greys") +
theme(plot.background=element_rect(fill="white"),
panel.background=element_rect(fill="white"),
axis.text=element_text(colour="black"),
text=element_text(size=20),
legend.position="bottom",
legend.title=element_blank()) +
### This plots another graph over the top
geom_bar(aes(fill=genotype), stat="identity",
colour=c("black"), # with black outline
show_guide=FALSE) # and no legend (the legend from the above code remains)
### ggplot2 for some reason puts a black slash through the boxes in the legend if the
black outline is specified. I further edited this figure in powerpoint to add simple
black boxes around the legend colors.

```



#Figure B.42

B.5 Statistical Analysis and Figures Presented in Chapter 4

The following will contain the R code used to analyze the data and create the figures presented in Chapter 4. Much of this is modified from code provided by Dr. Adam Davis.

B.5.1 Analysis of Dose Response Data and Creation of Dose Response Curves

```
### Importing the combined data from both runs and calling it "DRcomb"
DRcomb<-read.table("/Volumes/USB20FD/R/Dose Response - BCG/BCG-combined.txt",
header=T)

# Looking at the first few lines of the data
head(DRcomb)
#  INDEX Run REP Popn Dose DryWt PcontrolD PcontrolD1 PcontrolV PcontrolV1
#1  1  1  1  ACR  0.0  4.21 79.584121 0.79584121 100 1.0
#2  2  1  1  ACR  52.5  2.83 53.497164 0.53497165 80 0.8
#3  3  1  1  ACR 105.0  0.29 5.482042 0.05482042 0 0.0
#4  4  1  1  ACR 210.0  0.56 10.586011 0.10586011 40 0.4
#5  5  1  1  ACR 420.0  0.29 5.482042 0.05482042 0 0.0
#6  6  1  1  ACR 840.0  0.12 2.268431 0.02268431 0 0.0

### Determining whether the runs can actually be combined

### Running a two-way anova (and naming the model "m1") to determine whether
population, run or the interaction effect between them have a significant effect on
the percent control data (based on dry weight)
m1<-lm(DryWt~Popn+Run+ Popn:Run, data=DRcomb)

### View the results
summary.aov(m1)
#           Df Sum Sq Mean Sq F value  Pr(>F)
#Popn      4    354   88.53  12.535 8.03e-10 ***
```

```

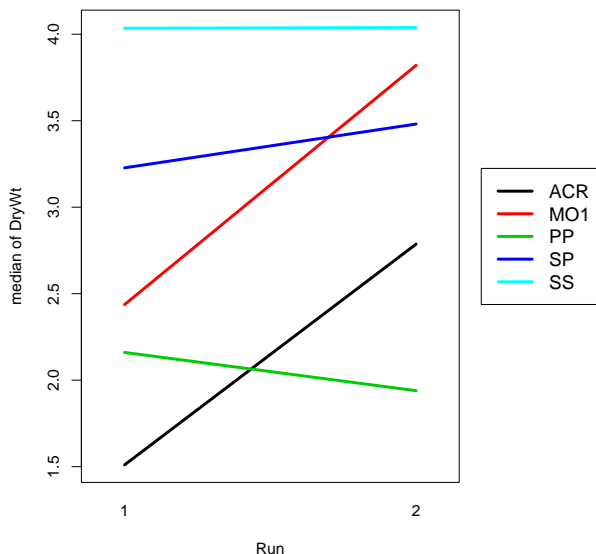
#Run      1      25  25.20  3.568  0.0594 .
#Popn:Run  4      53  13.34  1.888  0.1108
#Residuals 636  4492   7.06
#---
#Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

### There is no significant interaction effect between run and population - this means
that the populations are behaving similarly across runs and the data from runs can be
combined

### Making an interaction plot to examine how each population behaves from run to run
par(mar=c(5,5,5,10)) # setting the margins so I have room to put a legend on the right
side
interaction.plot(DRcomb$Run, DRcomb$Popn, DRcomb$DryWt,
  col=c(1:5), # specify colors (5 default colors)
  lty=1, lwd=3, # line type and line width
  legend=FALSE, # prevents legend from being made, it can be made later with the
legend command
  ylab="median of DryWt", # labeling the y axis
  xlab="Run", # labeling the x axis
  cex=1.2) # character expansion

### Adding the legend
locator(1) # this allows you to click anywhere in the graphing window once - then R
returns 1set of x,y coordinates. locator() allows unlimited clicking, press esc to
quit
#$x
#[1] 2.223273
#
#$y
#[1] 3.223943
par(xpd=TRUE) # allows plotting outside the plot area
legend(2.223273,3.223942, # plotting the legend where just clicked
legend=c("ACR", "MO1", "PP", "SP", "SS"), # text to go in the legend
  lty=1, lwd=3, col=c(1:5), cex=1.2) # same parameters as above in plot so legend
matches

```



#Figure B.43

```

### Installing and loading the drc package for analyzing dose responses
install.packages("drc") # This only needs to be done once per R program
library(drc)

```

```

### Creating the dose response model - four parameter log logistic model
drcomb.m.dry<-drm(PcontrolD~Dose, Popn, data=DRcomb,
fct=LL.4(names=c("b","lower","upper","ed50")))

### Viewing a summary of the model
summary(drcomb.m.dry)
#Model fitted: Log-logistic (ED50 as parameter) (4 parms)
#
#Parameter estimates:
#
#           Estimate Std. Error  t-value p-value
#b:ACR      1.93448    0.47064   4.11035  0.0000
#b:MO1      1.82483    0.45289   4.02931  0.0001
#b:PP       1.61304    0.37223   4.33343  0.0000
#b:SP       1.89887    0.26046   7.29045  0.0000
#b:SS       2.41384    0.37282   6.47458  0.0000
#lower:ACR  5.80290    3.88708   1.49287  0.1360
#lower:MO1  7.98860    5.87935   1.35876  0.1747
#lower:PP   5.18688    3.28588   1.57853  0.1149
#lower:SP   8.00204    2.90527   2.75432  0.0061
#lower:SS  13.93422    3.64134   3.82667  0.0001
#upper:ACR 100.06113    5.57559  17.94627  0.0000
#upper:MO1 100.78784    5.16477  19.51450  0.0000
#upper:PP  97.57886    6.06948  16.07696  0.0000
#upper:SP 105.21750    2.81003  37.44348  0.0000
#upper:SS 111.43477    2.74212  40.63819  0.0000
#ed50:ACR  72.59118    9.20860   7.88297  0.0000
#ed50:MO1 660.42120   98.83000   6.68240  0.0000
#ed50:PP  138.04746   21.56811   6.40054  0.0000
#ed50:SP  327.90975   29.18373  11.23605  0.0000
#ed50:SS  481.12970   44.25023  10.87293  0.0000
#
#Residual standard error:
#
#19.38802 (626 degrees of freedom)

### Doing the same thing for the visual data
drcomb.m.vis<-drm(DRcomb$PcontrolV~DRcomb$Dose, DRcomb$Popn, data=DRcomb,
fct=LL.4(names=c("b","lower","upper","ed50")))
summary(drcomb.m.vis)

### Calculating the ED values (Effective dose) for 10%, 50%, or 90% reduction in
growth
ED(drcomb.m.dry, c(10,50,90),interval="delta")
ED(drcomb.m.vis, c(10,50,90),interval="delta")
### Results look like . . .
#
#Estimated effective doses
#(Delta method-based confidence interval(s))
#
#           Estimate Std. Error  Lower  Upper
#ACR:10    23.3132     7.1588   9.2549  37.371
#ACR:50    72.5912     9.2086  54.5077  90.675
#ACR:90   226.0300    68.0386  92.4184 359.642
#MO1:10   198.1072    58.3011  83.6178 312.596
#MO1:50   660.4212   98.8300 466.3427 854.500

### More info
?ED

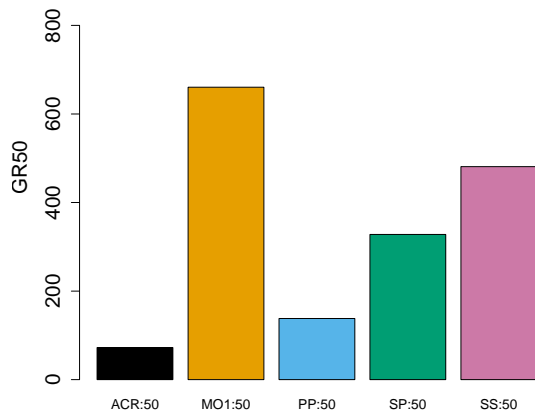
### Saving just the ED50 values - naming them "ED50dry" and "ED50vis"
ED50dry<-ED(drcomb.m.dry, c(50),interval="delta",level=0.95) #level is the confidence
interval level

```

```
ED50vis<-ED(drcomb.m.vis, c(50),interval="delta",level=0.95)
```

```
### Dry Plot: named EDd
EDd<-barplot(ED50dry[,1], # plotting all rows of column 1 from the ED50dry dataset
generated previously
col=c("#000000", "#E69F00", "#56B4E9", "#009E73", "#CC79A7"), # these specific colors
were chosen to be colorblind friendly (used for a seminar presentation)
ylab="GR50", # labeling the y axis
cex.lab=1.5, cex.main=1.75, cex.axis=1.5, # character expansion
ylim=c(0,max(ED50dry[,1])+300), # setting the range for the y axis (depends on the
ED50values), the maximum value will be the maximum value present in the first column
of the ED50dry data + 300 - This "300" can be increased or decreased to better to fit
the data if needed
main="Dry Weight GR50 Values") # assigning a main title
```

Dry Weight GR50 Values

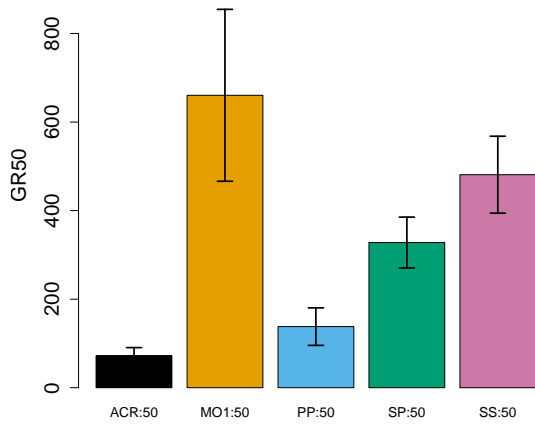


#Figure B.44

```
### Adding Error bars
```

```
segments(EDd, ED50dry[,3], EDd, ED50dry[,4], lwd=2) # draws a vertical line on the
graph("EDd") based on columns 3 (lower) and 4 (upper) from the ED50dry data
segments(EDd-0.1, ED50dry[,3], EDd+0.1, ED50dry[,3], lwd=2 ) # draws the bottom
horizontal hash
segments(EDd-0.1, ED50dry[,4], EDd+0.1, ED50dry[,4], lwd=2 ) # draws the upper
horizontal hash
```

Dry Weight GR50 Values

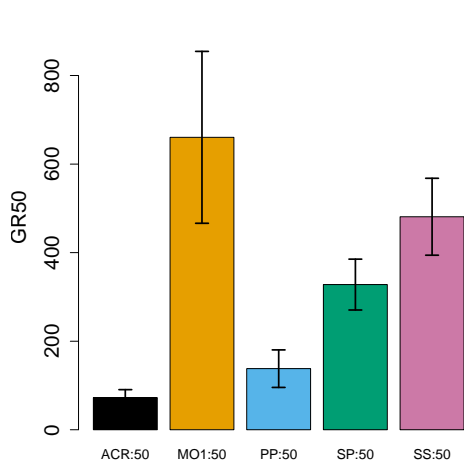


#Figure B.45

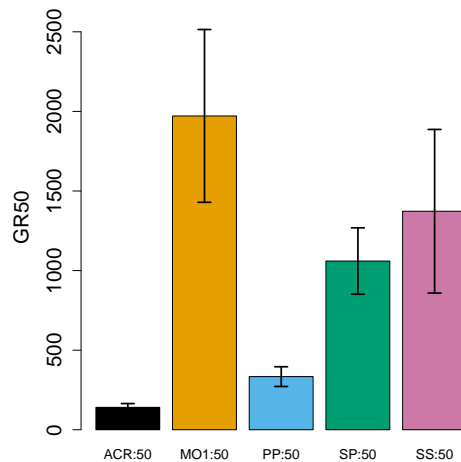
Putting it all together - doing the same thing with the visual data and plotting the dry and visual side by side

```
par(mfcol = c(1, 2)) # makes a graphing window with spots for 2 graphs so the visual
and dry graphs can go side by side in 1 window
### dry weight plot (same as above)
EDd<-barplot(ED50dry[,1], col=c("#000000","#E69F00","#56B4E9","#009E73", "#CC79A7"),
ylab="GR50", cex.lab=1.5, cex.main=1.75, cex.axis=1.5, ylim=c(0,max(ED50dry[,1])+300),
main="Dry Weight GR50 Values")
segments(EDd, ED50dry[,3], EDd, ED50dry[,4], lwd=2)
segments(EDd-0.1, ED50dry[,3], EDd+0.1, ED50dry[,3], lwd=2 )
segments(EDd-0.1, ED50dry[,4], EDd+0.1, ED50dry[,4], lwd=2 )
# visual data plot
EDv<-barplot(ED50vis[,1], col=c("#000000","#E69F00","#56B4E9","#009E73", "#CC79A7"),
ylab="GR50", cex.lab=1.5, cex.main=1.75, cex.axis=1.5, ylim=c(0,max(ED50vis[,1])+700),
main="Dry Weight GR50 Values")
segments(EDv, ED50vis[,3], EDv, ED50vis[,4], lwd=2)
segments(EDv-0.1, ED50vis[,3], EDv+0.1, ED50vis[,3], lwd=2 )
segments(EDv-0.1, ED50vis[,4], EDv+0.1, ED50vis[,4], lwd=2 )
```

Dry Weight GR50 Values



Dry Weight GR50 Values



#Figure B.46

```

### Comparing resistance levels between lines

### Computing the comparison and naming it "DryParm"
DryParm<-compParm(drcomb.m.dry, "ed50")
### Finding the reciprocal to obtain a number for every comparison
1/DryParm[,1]

### The new std errors are not calculated by taking the reciprocal, to get the correct
standard error, perform a manual comparison
# First get the ED50 values (naming them as an object called "ED50s")
ED50s<-ED(drcomb.m.dry, c(50),interval="delta")
#Estimated effective doses
#(Delta method-based confidence interval(s))
#
#      Estimate Std. Error   Lower   Upper
#ACR:50  72.5912     9.2086  54.5077  90.675
#MO1:50 660.4212    98.8300 466.3427 854.500
#PP:50  138.0475    21.5681  95.6929 180.402
#SP:50  327.9098    29.1837 270.5999 385.220
#SS:50  481.1297    44.2502 394.2328 568.027

### Perform a manual comparison using the ED50 values
# MO1/ACR
comped(c(660.4212,72.5912),c(98.8300,54.5077), log=FALSE, operator="/")
# First set of numbers is the Estimate, second set is the Std. Errors
# Example: comped(c(MO1 estimate, ACR estimate), c(MO1 Std. Error, ACR Std. Error),
log=FALSE, operator="/")
#Estimated ratio of effective doses
#
#      Estimate Std. Error   Lower Upper
#[1,]   9.0978     6.9658 -4.5548 22.75

### Calculating R/S ratios using ACR as the sensitive, but using indexing to call
numbers from the table of ED50 values instead of typing in the actual numbers
# ACR/ACR
comped(c(ED50s[1,1],ED50s[1,1]),c(ED50s[1,2],ED50s[1,2]), log=FALSE, operator="/")
# PP/ACR
comped(c(ED50s[3,1],ED50s[1,1]),c(ED50s[3,2],ED50s[1,2]), log=FALSE, operator="/")
# SP/ACR
comped(c(ED50s[4,1],ED50s[1,1]),c(ED50s[4,2],ED50s[1,2]), log=FALSE, operator="/")
# SS/ACR
comped(c(ED50s[5,1],ED50s[1,1]),c(ED50s[5,2],ED50s[1,2]), log=FALSE, operator="/")

### Calculating R/S ratios using PP as the sensitive
# ACR/PP
comped(c(ED50s[1,1],ED50s[3,1]),c(ED50s[1,2],ED50s[3,2]), log=FALSE, operator="/")
# MO1/PP
comped(c(ED50s[2,1],ED50s[3,1]),c(ED50s[2,2],ED50s[3,2]), log=FALSE, operator="/")
# PP/PP
comped(c(ED50s[3,1],ED50s[3,1]),c(ED50s[3,2],ED50s[3,2]), log=FALSE, operator="/")
# SP/PP
comped(c(ED50s[4,1],ED50s[3,1]),c(ED50s[4,2],ED50s[3,2]), log=FALSE, operator="/")
# SS/PP
comped(c(ED50s[5,1],ED50s[3,1]),c(ED50s[5,2],ED50s[3,2]), log=FALSE, operator="/")

### Same thing as above but with the visual data (no manual comparisons shown)
VisParm<-compParm(drcomb.m.vis, "ed50")
1/VisParm[,1]

### Plotting the Data - there are many ways, several examples follow

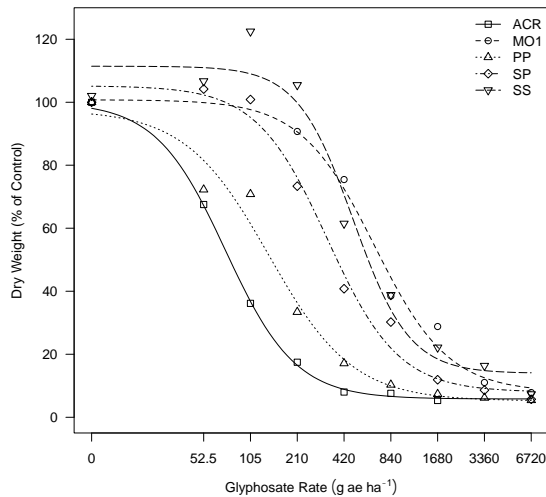
```



```

### Example 1
plot(drcomb.m.dry, # plotting the model created above called drcomb.m.dry
broken=FALSE, # no break in the x axis
xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})), # labeling the x axis
type="average", # will plot the averages
ylab="Dry Weight (% of Control)", # labeling the y axis
ylim=c(0,125), # setting the ranges of the y axis
col=c("black"), # color of lines
lty=c(1:6), # specifying line type
xt=c(0,52.5,105,210,420,840,1680,3360,6720), #specifying the x axis values with rates
used
xtrim="FALSE", # don't trim the x axis values
legend=FALSE, # no legend (will be made later)
pch=c(0,1,2,5,6) # plotting character
legend("topright", # adding a legend in the top right corner of the plot area
legend=c("ACR", "MO1", "PP", "SP", "SS"), # text for legend
lty=c(1:6),pch=c(0,1,2,5,6), # same parameters as used for making the figure
bty="n") # no box drawn around legend

```

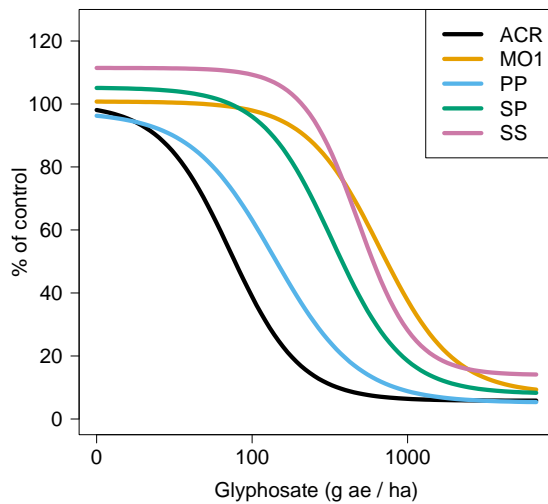


#Figure B.47

```

### Example 2
plot(drcomb.m.dry, broken=FALSE, xlab = "Glyphosate (g ae / ha)", ylab = "% of
control", type=c("none"), # will not plot any points - just lines
ylim=c(0,125),
col=c("#000000", "#E69F00", "#56B4E9", "#009E73", "#CC79A7"), # colorblind friendly color
scheme
lty=(1), lwd=5, legend=FALSE, cex=1.5, cex.axis=1.5, cex.main=1.75, cex.lab=1.5)
legend("topright", legend=c("ACR", "MO1", "PP", "SP", "SS"),
lty=1, lwd=5, col=c("#000000", "#E69F00", "#56B4E9", "#009E73", "#CC79A7"), cex=1.5)
# matching legend

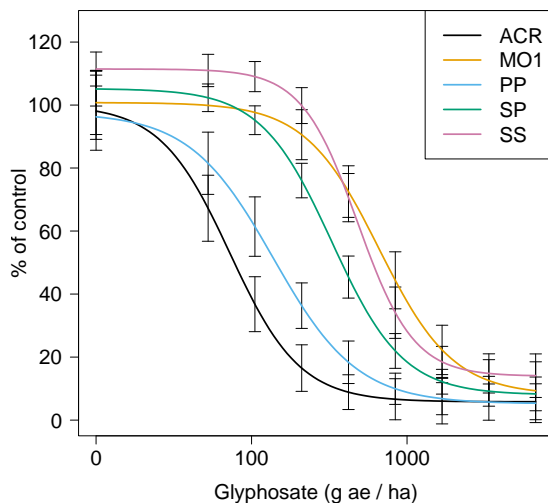
```



#Figure B.48

Example 3

```
plot(drcomb.m.dry, broken=FALSE, xlab = "Glyphosate (g ae / ha)", ylab = "% of control", type=c("bars"), # plots error bars
ylim=c(0,125), col=c("#000000","#E69F00","#56B4E9","#009E73", "#CC79A7"), # colorblind friendly color scheme
lty=(1), lwd=2, legend=FALSE, cex=1.5, cex.axis=1.5,cex.main=1.75, cex.lab=1.5)
legend("topright", legend=c("ACR", "MO1", "PP", "SP", "SS"),
lty=1, lwd=2, col=c("#000000","#E69F00","#56B4E9","#009E73", "#CC79A7"), cex=1.5)
```



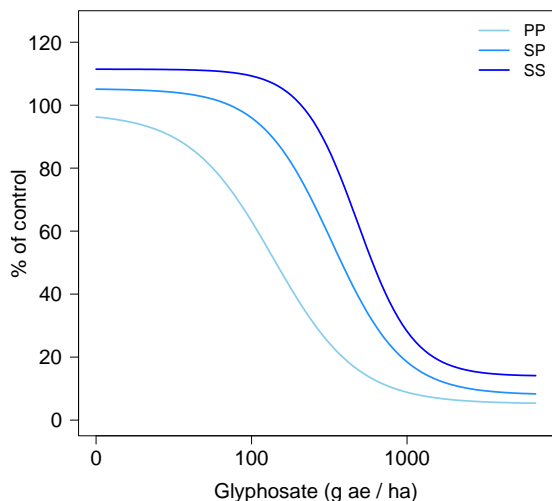
#Figure B.49

```
### the levels=c("") command allows you to plot specific lines
### the type=c("average", "obs", "none", "all", "bars") command changes graph type
##### average = default
##### obs = just plots the observations
##### none = plots just the curves
##### all = plots curves and all observations
##### bars = plots error bars
```

```

### Example 4 - Only the BCG genotypes
plot(drcomb.m.dry, broken=FALSE,
level=c("PP", "SP", "SS"),
xlab = "Glyphosate (g ae / ha)", ylab = "% of control", type="none",
ylim=c(0,125), col=c("skyblue", "dodgerblue", "blue"),
lty=(1), lwd=2, legend=FALSE, cex=1.2, cex.axis=1.5, cex.main=1.75, cex.lab=1.5)
legend("topright", legend=c("PP", "SP", "SS"),
      lty=1, lwd=2, col=c("skyblue", "dodgerblue", "blue"), cex=1.2, bty="n")

```



#Figure B.50

```

### More info:
?plot.drc

```

The dose response performed on sensitive controls was analyzed very similarly to the BCG dose response above and therefore is minimally annotated

```

### Importing data

```

```

DR<-read.table("/Volumes/USB20FD/R/Dose Response - BCG/ACR_MCR_WCS.txt", header=T)
head(DR)

```

```

# Popn Rep Dose Vis Vis.1 Weight PcontrolD PcontrolDl PcontrolV PcontrolVl
#1 ACR 1 0.0 0 100 5.94 1.00000000 100.000000 1.0 100
#2 ACR 1 52.5 50 50 1.29 0.21717172 21.717172 0.5 50
#3 ACR 1 105.0 100 0 0.59 0.09932660 9.932660 0.0 0
#4 ACR 1 210.0 70 30 0.74 0.12457912 12.457912 0.3 30
#5 ACR 1 420.0 100 0 0.35 0.05892256 5.892256 0.0 0
#6 ACR 1 840.0 100 0 0.39 0.06565657 6.565657 0.0 0
# PcontrolDA PcontrolDA1 PcontrolVA PcontrolVA1
#1 1.02149613 102.149613 1.0344828 103.44828
#2 0.22184007 22.184007 0.5172414 51.72414
#3 0.10146174 10.146174 0.0000000 0.000000
#4 0.12725709 12.725709 0.3103448 31.03448
#5 0.06018917 6.018917 0.0000000 0.000000
#6 0.06706793 6.706793 0.0000000 0.000000

```

```

### Computing dose response models (4 parameter log logistic)

```

```

drdry<-drm(DR$PcontrolDl~DR$Dose, DR$Popn, data=DR,
fct=LL.4(names=c("b", "lower", "upper", "ed50")))
summary(drdry)

```

```

drvis<-drm(DR$PcontrolVl~DR$Dose, DR$Popn, data=DR,
fct=LL.4(names=c("b", "lower", "upper", "ed50")))
summary(drvis)

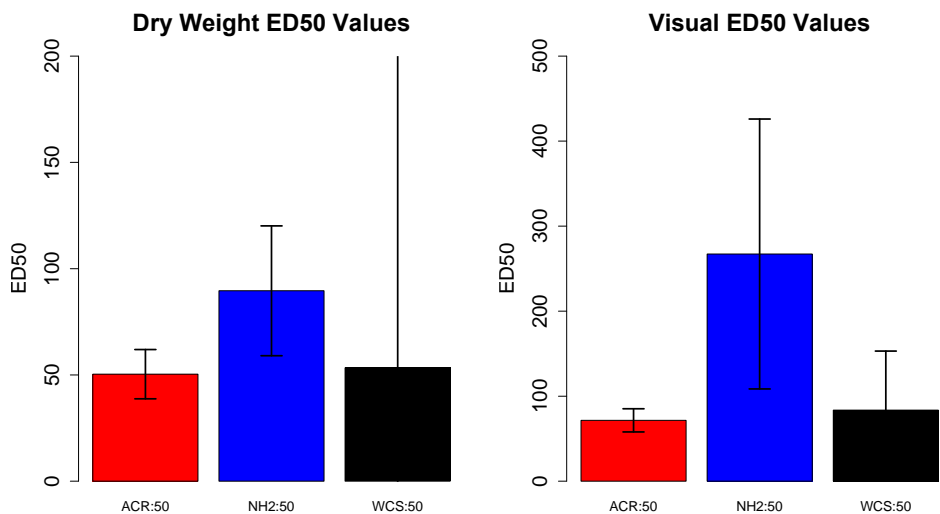
```

```

### Computing ED values and Plotting
ED(drdry, c(10,50,90),interval="delta", level=0.95)
ED(drvis, c(10,50,90),interval="delta")

par(mfcol = c(1, 2))
ED50dry<-ED(drdry, c(50),interval="delta",level=0.95)
ED<-barplot(ED50dry[,1], col=c("Red", "Blue", "Black"), ylab="ED50",
cex.lab=1.5, cex.main=1.75, cex.axis=1.5, ylim=c(0,200),
main="Dry Weight ED50 Values")
segments(ED, ED50dry[,3], ED, ED50dry[,4], lwd=2)
segments(ED+0.1, ED50dry[,3], ED+0.1, ED50dry[,3], lwd=2 )
segments(ED+0.1, ED50dry[,4], ED+0.1, ED50dry[,4], lwd=2 )
ED50vis<-ED(drvis, c(50),interval="delta",level=0.95)
EDv<-barplot(ED50vis[,1], col=c("Red", "Blue", "Black"), ylab="ED50",
cex.lab=1.5, cex.main=1.75, cex.axis=1.5, ylim=c(0,500),
main="Visual ED50 Values")
segments(EDv, ED50vis[,3], EDv, ED50vis[,4], lwd=2)
segments(EDv+0.1, ED50vis[,3], EDv+0.1, ED50vis[,3], lwd=2 )
segments(EDv+0.1, ED50vis[,4], EDv+0.1, ED50vis[,4], lwd=2 )

```



#Figure B.51

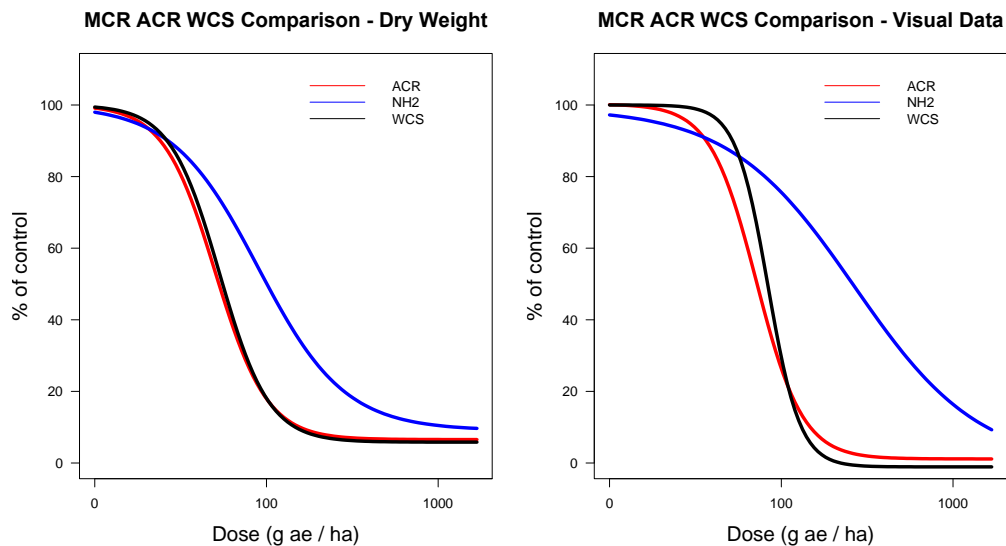
```

### Finding R/S values
dED50s<-ED(drdry, c(50),interval="delta")
drycomp<-compParm(drdry, "ed50")
1/drycomps[,1]

### Manual comparison to get correct std. errors
comped(c(dED50s[1,1],dED50s[1,1]),c(dED50s[1,2],dED50s[1,2]), log=FALSE, operator="/")
comped(c(dED50s[2,1],dED50s[1,1]),c(dED50s[2,2],dED50s[1,2]), log=FALSE, operator="/")

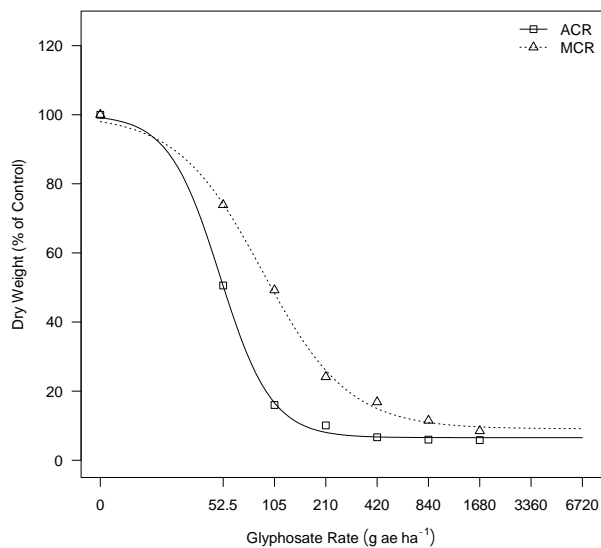
### Plotting the data
par(mfcol = c(1, 2))
plot(drdry, broken=FALSE,
main = "MCR ACR WCS Comparison - Dry Weight",
xlab = "Dose (g ae / ha)", ylab = "% of control", type=c("none"),
col =c("Red", "Blue", "Black"), lty=(1), pch=(15:19), ylim=c(0,110), lwd=4,
cex.main=1.5, cex.lab=1.5)
plot(drvis, broken=FALSE,
main = "MCR ACR WCS Comparison - Visual Data",
xlab = "Dose (g ae / ha)", ylab = "% of control", type=c("none"),
col =c("Red", "Blue", "Black"), lty=(1), pch=(15:19), ylim=c(0,110), lwd=4,
cex.main=1.5, cex.lab=1.5)

```



#Figure B.52

```
### Plotting only MCR and ACR for Chapter 4
plot(drdry, level = c("ACR", "NH2"), broken=FALSE,
xlab=expression("Glyphosate Rate"~(g ~ ae ~ ha^{-1})), type="average",
ylab="Dry Weight (% of Control)",
ylim=c(0,125), col=c("black"), lty=c(1,3),
xt=c(0,52.5,105,210,420,840,1680,3360,6720), xtrim="FALSE", xlim=c(0,6720), legend =
FALSE, pch=c(0,2))
legend("topright", legend=c("ACR", "MCR"), lty=c(1,3), bty="n", pch=c(0,2))
```



#Figure B.53

B.6 Literature Cited

Adler D (2005) vioplot: Violing plot. R package version 0.2. <http://wsopuppenkiste.wiso.uni-goettingen.de/~dadler>

Becker AR, Wilks AR, Brownrigg R, Minka TP (2013). maps: Draw Geographical Maps. R package version 2.3-6. <http://CRAN.R-project.org/package=maps>

Neuwirth E (2011) RColorBrewer: ColorBrewer palettes. R package version 1.0-5. <http://CRAN.R-project.org/package=RColorBrewer>

R Development Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <http://www.R-project.org> Accessed March 26, 2014

Ritz C, Streibig JC (2005) Bioassay Analysis using R. J. Statist. Software, 12:5.

Wheeler B (2010) lmPerm: Permutation tests for linear models. R package version 1.1-2. <http://CRAN.R-project.org/package=lmPerm>

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer New York. <http://had.co.nz/ggplot2/book>