

BIOCHEMICAL AND MOLECULAR BASIS FOR METABOLISM-BASED
ATRAZINE RESISTANCE IN *AMARANTHUS TUBERCULATUS*

BY

ANTON FREDRICK EVANS JR.

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 Dean E. Riechers
Associate Professor Kris N. Lambert
Professor Steven C. Huber
Professor Patrick J. Tranel

Abstract

Atrazine, a photosystem II inhibitor, is a commonly used herbicide for selective broadleaf weed control in maize (*Zea mays*). Rapid detoxification of s-triazine herbicides is the result of increased glutathione S-transferase (GST) activity in atrazine-tolerant crops such as maize and grain sorghum (*Sorghum bicolor*), or in atrazine-tolerant weeds such as *Panicum* spp. In previous research, two atrazine-resistant populations of tall waterhemp (*Amaranthus tuberculatus*) from central Illinois (designated ACR and MCR) did not possess a point mutation in *psbA* (encoding the D1 protein, the target site for atrazine), but rapid production of atrazine-glutathione conjugates indicated that elevated rates of metabolism may be a major resistance mechanism. My objective was to utilize traditional protein purification techniques combined with proteomic methods to further investigate the hypothesis that enhanced metabolic detoxification of atrazine, catalyzed by distinct GST isozymes, may confer resistance in ACR and MCR. ACR and MCR showed higher GST specific activities towards atrazine in partially-purified ammonium sulfate and glutathione affinity-purified fractions when compared with an atrazine-sensitive waterhemp population (WCS). One-dimensional gel electrophoresis displayed a large band at approximately 26 kDa, which is the typical subunit mass of GST subunits. Identification of several GSTs by LC-MS/MS of glutathione affinity-purified fractions from each population led to the identification of a phi class GST (*ArGSTF1*) and two tau class GSTs, based on peptide sequence similarity with GSTs from *Arabidopsis*. Elevated constitutive expression levels of *ArGSTF1* in ACR and MCR correlated with atrazine resistance, while expression of each tau class GST was highest in MCR but low in ACR and WCS.

Acknowledgments

I would like to thank Dr. Dean Riechers for acting as my adviser, for his knowledge, guidance and providing me with the opportunity to work with him. I want to thank all the members of Riechers Lab including Rong Ma, Josh Skelton, Dr. Anatoli Lygin, Brittany Janney, and Jacqueline Janney. I would also like to thank my committee members Dr. Kris Lambert, Dr. Steven Huber, and Dr. Patrick Tranel for all the help and support they gave me. Finally I would like to thank Dr. Isaac Cann for giving my first opportunity to pursue my dream of becoming a scientist.

Table of Contents

LITERATURE REVIEW.....	1
INTRODUCTION.....	9
MATERIALS AND METHODS.....	14
RESULTS.....	22
CONCLUSIONS.....	28
REFERENCES.....	32
TABLES AND FIGURES.....	39
APPENDIX.....	44

Literature Review

The waterhemp problem

In the United States, management of waterhemp [*Amaranthus rudis* or *Amaranthus tuberculatus* (Pratt and Clark, 2001)] has become problematic in maize (*Zea mays*) and soybean (*Glycine max*) cropping systems. Worldwide there are about 75 species belonging to the *Amaranthus* genus with both monoecious and dioecious species (Pratt and Clark, 2001; Steckel, 2007). Waterhemp is one of the 10 dioecious species of amaranth found in North America. Waterhemp is a dicot with a C₄ photosynthetic system (Steckel, 2007), and is a summer annual with germination and emergence periods extending late into the growing season (Hager et al., 2002). This, along with the decreased use of soil-applied residual herbicides (Sprague et al., 1997) increased adoption of reduced tillage systems, and the rapid rise of herbicide-resistant biotypes (Steckel, 2004; Heap, 2014), has resulted in waterhemp becoming a major problem in corn and soybean production. For example, waterhemp competition can reduce soybean and corn yields by as much as 56% and 74% respectively (Steckel, 2007). Waterhemp populations have been identified that are resistant to 4-hydroxyphenylpyruvate dioxygenase inhibitors, acetolactate synthase inhibitors, protoporphyrinogen oxidase inhibitors, glyphosate, 2,4-D, and photosystem II inhibitors (Hausman et al., 2011; Heap 2014).

Photosystem II inhibitors

Photosystem II (PSII) is one of two photochemical reaction centers found in plants (Buchanan, 2000). The D1 and D2 proteins and chlorophyll P680 are the major components of this complex. In normal reactions light energy is absorbed by P680. The excited electrons are then transferred to two plastoquinone molecules at the Q_a and Q_b binding sites within the D2 and D1 proteins, respectively. P680 is reduced by electrons provided by the D1 protein, and reduced plastoquinone then passes the electrons to the cytochrome b_6f complex (Buchanan, 2000; Hess, 2000).

Herbicides that inhibit the PSII mechanism competitively inhibit the binding of plastoquinone to the Q_b site of the D1 protein. This results in the excitation of chlorophyll to a triplet state. Triplet chlorophyll, which is unstable, then induces rapid formation of singlet oxygen. Lipid peroxidation follows as a result of both triplet chlorophyll and singlet oxygen inducing lipid and hydroxyl radicals, and cell death quickly follows (Hess, 2000). Chemicals that inhibit PSII belong to many sub-classes such as dinitrophenols, biscarbamates, phenyl ureas, nitriles, triazinones, and symmetrical (s)-triazines (Oettmeier, 1992).

Atrazine, an s-triazine, is one of the most commonly used herbicides in maize in the United States, with an estimated use on 65-70% of all the corn acreage (LeBaron et al., 2011). Atrazine can be applied either postemergence or preemergence to the soil, which provides residual control on sensitive dicot weeds up to 60 days depending on soil conditions (Krutz et al., 2009). Atrazine injury is typically characterized by yellowing of the older leaves followed by necrosis starting around the leaf margins. This is due to translocation of atrazine through the xylem only (Lebaron et al., 2011). Resistance to

atrazine and other PSII inhibitors has arisen from mutations at the site of action and as a result of enhanced metabolism (Yuan et al., 2007, Nemat Alla et al., 2008, Lebaron et al., 2011)

Target site based resistance to PSII inhibitors

Target site-based resistance to atrazine and other PSII inhibitors has been documented in several weeds including several *Amaranthus* species (Hirschberg and McIntosh, 1983; Heap, 2014). Several mutations in the *psbA* gene, which encodes the D1 protein, can affect the Q_b binding site. This allows plastoquinone to continue to bind, but blocking the binding of the herbicide. Resistance to the highest use rates of the particular herbicide results, but at the cost of general productivity due to less efficient binding of plastoquinone, which results in less efficient electron transport (Holt et al., 1993; Lebaron et al., 2011). Several amino acid substitutions can result in an insensitive target site and confer weed resistance to atrazine. These substitutions include Phe-211 to Ser (Gingrich et al., 1988), Ala-251 to Val, Phe-256 to Tyr, Ser-264 to Ala (Lebaron et al., 2011), and the most common substitution found in *Amaranthus* and other weeds, Ser-264 to Gly (Hirschberg and McIntosh, 1983).

Metabolism based resistance to PSII inhibitors

Resistance to PSII inhibitors can also arise from enhanced herbicide detoxification due to metabolism. Tolerance to atrazine and other PSII herbicides in maize, grain sorghum, and wheat results from rapid detoxification mechanisms (Nemat

Alla et al., 2008). Metabolism-based resistance occurs when the parent herbicide is altered by plant enzymes, rendering it more water soluble, immobile, and non-toxic. Metabolism has to occur at a rate high enough to prevent significant herbicide binding to the site of action. In the case of PSII inhibitors, detoxification has to occur faster than significant binding of the Q_b site of the D1 protein. There are several classes of plant enzymes placed into three phases known for their distinct roles in detoxification reactions in plants.

Phase I is characterized by oxidative enzymes such as cytochrome P450s. P450s are found in all domains of life and are generally characterized by their ability to insert an oxygen atom to an aliphatic position of an organic substrate in a monooxygenase reaction that requires molecular oxygen, NADPH, and electrons from cytochrome P450 reductases (Buchanan, 2000). P450s are also involved in dealkylation, deamination, decarboxylation, and other oxidative reactions that result in more water soluble or reactive products that are ready for Phase II detoxification (Riechers et al., 2010).

Glycosyltransferases (uGTs) are another class of enzymes typical of Phase II detoxification reactions. uGTs are characterized by their ability to conjugate xenobiotics with a sugar molecule. This glycosylation reaction can occur with hydroxyl, carboxyl, sulfhydryl, and amino functional groups (Bowles et al., 2005). In addition to their roles in detoxification, uGTs are involved in maintaining plant hormone concentrations (Yuan et al., 2007). These sugar-conjugated molecules are then targeted for Phase III detoxification.

ATP-binding cassette (ABC) transporters are the major proteins involved in the process of Phase III detoxification in the tonoplast membrane (Riechers et al. 2010). ABC transporters use energy derived from the hydrolysis of ATP to actively transport a variety of conjugated substrates, and are involved in the vacuolar transport of heavy metals, fatty acids, auxins, phospholipids, and glutathione- or glycosylated herbicide conjugates (Schulz and Kolukisaoglu, 2006; Riechers et al., 2010). These compounds are then targeted for cell wall or vacuolar sequestration. For example, glucoside-conjugated hydroxy-primisulfuron is targeted to the vacuole in *Hordeum vulgare* (Gaillard et al., 1994). In this detoxification pathway, a parent herbicide requires modification by a P450 in Phase I to become more reactive or water soluble, but also requires conjugation in Phase II in order to completely detoxify the herbicide since some hydroxylated or de-alkylated herbicides remain partially phytotoxic or phloem mobile (Riechers et al., 2010). In Phase III, the conjugated compound is sequestered in the vacuole or deposited in cell walls as bound residues through a process mediated by ABC transporters. However, there is an additional class of Phase II detoxification enzymes that can directly conjugate parent herbicides without prior modification by P450s in Phase I, which are described below.

Glutathione S-transferases (GSTs) are found in all eukaryotic cells and are well-studied due to their ability to detoxify a wide variety of xenobiotic and endogenous substrates (Riechers et al., 2010). GSTs are characterized by catalyzing the conjugation of reduced glutathione (GSH), (2S)-2-amino-4-[[[(1R)-1-[(carboxymethyl) carbamoyl]-2-sulfanylethyl]carbamoyl]butanoic acid), with a xenobiotic substrate (such as an herbicide) via nucleophilic displacement. The GSH-conjugated compound is then

rendered more water soluble, immobile, and non-toxic, and can then be targeted for Phase III detoxification (via tonoplast ABC transporters) to the vacuole for long-term storage, catabolism to cysteine conjugates, or degradation (Edwards and Dixon, 2005; Riechers et al., 2010).

Plant GST isozymes are encoded by multi-gene families and are categorized into several protein classes including theta, zeta, lambda, phi, tau, and GSH-dependent dehydroascorbate reductases (DHARs), which are based on their amino acid sequences, catalytic residues, immunological cross-reactivity, and genomic structures (Smith et al., 2004; Frova, 2006). Theta, lambda, and DHAR-class GSTs are the most recent plant-specific classifications and are limited to a few *Arabidopsis* genes with rice and soybean orthologs (Dixon et al., 2002). Zeta class GSTs are typically induced during senescence by ethylene. Tau class GSTs are induced by many stress factors including extreme temperatures, pathogens, heavy metal toxicity, natural and synthetic auxins, and herbicide safeners (Wagner et al., 2002; Frova, 2003; Edwards and Dixon, 2005). Phi class GSTs were some of the first discovered and are known for their herbicide detoxification activity (Frova, 2003).

Phi and tau class GSTs are the most abundant in plants, and are well known for their ability to conjugate a wide variety of herbicides with GSH, leading to herbicide selectivity in crops or weed resistance (Edwards and Dixon, 2005). Both GST enzyme classes are functionally active as both homodimers and heterodimers, and contain an essential catalytic serine residue near the N-terminus of the protein. Dimerization within a GST subclass can affect specificity and catalytic activity with certain herbicide substrates (Holt et al., 1995); however, cross dimerization between the tau and phi

classes has not been shown (Cummins et al., 2011). Tau and phi class GSTs are involved in herbicide metabolism and resistance in *Echinochola crus-galli* (barnyard grass), *Oryza sativa* (Yang et al., 2009), *Triticum tauschii* (Zhang et al., 2004) and *Arabidopsis* (Smith et al., 2004).

Atrazine tolerance in crops, such as *Zea maize* and grain sorghum, is a direct result of rapid metabolism due to GST activities (Timmerman, 1989; Nemat Alla et al., 2008; Li et al., 2013). GST-catalyzed herbicide metabolism is also a major contributing factor in conferring atrazine resistance in populations of *Abutilon theophrasti* (Gray et al., 1996). Utilizing rapid metabolism as a biochemical resistance mechanism to alleviate the herbicidal stress of atrazine allows these plants to avoid decreased productivity (*i.e.*, a fitness cost) that typically results from an altered D₁ target site (Holt et al., 1993; Nemat Alla et al., 2008). This mechanism is similar to what was observed in atrazine-resistant populations of waterhemp from Illinois (Ma et al., 2013)

Studying atrazine resistance in waterhemp

It was determined by previous research that due to the lack of a mutation in the *psbA* gene, atrazine-resistant populations of *Amaranthus tuberculatus* from McLean County and Adams County, Illinois (MCR and ACR, respectively) most likely resulted from non-target site mechanism(s). Subsequent metabolism studies showed rapid accumulation of a non-polar metabolite with the same retention time as a synthetic GSH-atrazine standard in MCR and ACR (Ma et al., 2013), indicating enhanced atrazine metabolism by GST activities in atrazine-resistant populations of waterhemp.

The objective of my thesis research was to determine if differences existed in GST activity and/or expression between the atrazine-resistant populations and WCS. To accomplish this I used several biochemical approaches, where I initially focused on measuring GST activity with atrazine *in vitro* in crude protein extracts from leaves of each *A. tuberculatus* population. After obtaining evidence of differential GST activity with atrazine in partially-purified GST fractions from atrazine-resistant and -sensitive plants, I then utilized proteomic and molecular approaches in which I identified amino acid sequences from the one-dimensional gel-purified GSTs, matched them with protein and mRNA sequences from *Arabidopsis* and *A. tuberculatus*, and finally measured constitutive transcript expression levels in atrazine-resistant and -sensitive populations via qRT-PCR and gene-specific primers. In addition, I screened several *Amaranthus* tubulin genes to determine the most suitable constitutively-expressed transcript to normalize my GST expression data.

Introduction

There are over seventy-five species in the genus *Amaranthus* found worldwide, including both monoecious and dioecious species (Steckel, 2007). A dioecious species called waterhemp (*Amaranthus tuberculatus* or *A. rudis*) (Pratt and Clark, 2001) has become a major problem in the central and southern U.S. (Steckel, 2007) due to several biological factors described below. Waterhemp is difficult to selectively control in maize (*Zea mays*) and soybean (*Glycine max*) production systems because it is a summer annual with a prolonged germination season that extends late into the summer, and the obligate out-crossing nature and dioecious biology of waterhemp has ensured that genes enabling herbicide resistance are rapidly spread throughout populations (Steckel, 2007, Hausman et al., 2011; Ma et al., 2013). Many of these genes have been stacked within single waterhemp populations and individual plants, resulting in multiple herbicide-resistant phenotypes (Heap 2014). For example, resistance to herbicides that inhibit 4-hydroxyphenylpyruvate dioxygenase, protoporphyrinogen oxidase, acetolactate synthase, photosystem II (PSII), and the synthetic auxin herbicide 2,4-D has been reported (Hausman et al., 2011; Ma et al., 2013).

Atrazine (1-chloro-3-ethylamino-5-isopropylamino-2,4,6-triazine) belongs to a group of herbicides that bind to the D₁ protein in PSII (LeBaron et al., 2011; Hess, 2000). These inhibitors block the flow of electron transport during photosynthesis. Atrazine inhibits PSII electron transport by competing with plastoquinone for the Q_b binding site of the D₁ protein. This results in electrons not being passed to cytochrome B₆f, and subsequently the rapid formation of triplet chlorophyll followed by singlet

oxygen in the presence of light. Singlet oxygen then interacts with the cell membranes, which results in lipid peroxidation, reactive oxygen species (ROS), and rapid cell death (Hess, 2000).

There are two major classes of PSII inhibitors that are still used for selective weed control, which include several compounds in the phenyl urea and triazine subclasses. Atrazine, a symmetrical triazine (s-triazine), is one of the most commonly used herbicides for weed management in maize, with an estimated use on 65-70% of all maize acreage in the U.S. (Lebaron et al., 2011). The most common resistance mechanism to atrazine is an insensitive target site. In atrazine-resistant weed biotypes, a point mutation in the *psbA* gene (which encodes the D₁ protein of PSII) results in a serine to glycine mutation at amino acid 264, which confers an approximate 1000-fold level of atrazine resistance compared with sensitive biotypes (Hirschberg and McIntosh, 1983; Holt et al., 1993; Gray et al., 1996). However, this mutation also results in a significant fitness penalty due to the lower affinity of the D₁ protein for plastoquinone binding (Holt et al., 1993; Gray et al., 1996) and a less efficient rate of electron transport in PSII, leading to reduced biomass production (Holt et al., 1993). Natural tolerance in maize and grain sorghum is due to the high activity of glutathione S-transferase (GSTs) that can use atrazine as a substrate, leading to rapid metabolic detoxification in these crops as well as in naturally tolerant weed species such as *Panicum* (Thompson, 1972). Evolved resistance to atrazine in velvetleaf (*Abutilon theophrasti*) has also been linked to increased GST activity (Gray et al., 1996; Woodyard et al., 2009). Rapid detoxification results in up to 100-fold resistance compared to sensitive plants.

However, fitness costs and reduced growth rates are typically not observed as a result of this herbicide resistance mechanism (Gray et al., 1996; Woodyard et al., 2009).

GSTs are found in both plant and animals and are a widely studied class of cytosolic enzymes due to their detoxification abilities (Timmerman, 1989, Edwards and Dixon, 2005; Frova, 2006). GSTs function by catalyzing the conjugation of reduced glutathione (GSH) with xenobiotics containing an electrophilic site, such as s-triazine and chloroacetamide herbicides (Riechers et al., 2010). Xenobiotic-GSH conjugates are then transported to the vacuole by tonoplast transporters, where they are further degraded by peptidase enzymes (Wolf et al. 1996; Edwards and Dixon, 2005). Additional functions and activities of plant GSTs have also been reported, such as glutathione peroxidase (GPOX), dehydroascorbate reductase (DHAR), and oxidoreductase (Edwards and Dixon, 2005).

Plant GSTs belong to several different classes, including theta, zeta, lambda, phi, tau, and GSH-dependent dehydroascorbate reductase (DHAR), based on sequence similarity, essential catalytic residues, and immunological cross-reactivity (Edwards and Dixon, 2005; Smith et al., 2004). They are often found in clusters of duplicated genes within a chromosome. Comparative genomic analysis of these regions in *Oryza sativa* and *Arabidopsis thaliana* suggest that these regions prone to rapid changes in sequence promoting gene evolution (Soranzo et al., 2004). The most abundant GSTs in plants belong to the phi and tau classes, which are well known for their ability to conjugate a wide variety of herbicide substrates with GSH in cereal crops (Edwards and Dixon, 2005; Frova, 2003, 2006). Phi class GSTs were among the first plant GSTs shown to be involved with herbicide detoxification reactions (Timmerman, 1989) and

recently a phi-class GST was identified that confers aphid resistance in wheat plants (*Triticum aestivum*) (Shultz et al. 2014, in press). Tau-class GSTs are frequently identified in gene expression studies in plants due to their induction by stress factors such as heavy metals, temperature, drought, wounding, phytohormones, allelochemicals, herbicides, and herbicide safeners (Wagner et al. 2002; Frova, 2003, 2006; Riechers et al., 2010). Both classes of GSTs are involved in herbicide tolerance in maize (Timmerman, 1989), and rice (*Oryza sativa*) (Yang et al., 2009) and herbicide resistance in barnyardgrass (*Echinochola crus-galli*) (Li et al., 2013), a diploid wheat species (*Triticum tauschii*) (Zhang et al., 2004), and *Arabidopsis* (Smith et al., 2004).

Tau and phi GSTs are functionally active as both homodimers and heterodimers, each with a conserved GSH binding domain in the N-terminus (Cummins et al., 2011; Edwards 2000). Although not yet thoroughly investigated, it has been hypothesized that heterodimers may demonstrate unique kinetic properties, uniquely recognize biologically active compounds within the plant, or regulate protein localization or turnover (Edwards et al., 2000). However, cross dimerization between subunits of the tau and phi GST subclasses has not been shown *in vitro* (Cummins et al., 2011).

Previous research demonstrated that atrazine resistance in two populations of waterhemp from Illinois results from non-target site resistance (NTSR) mechanism(s), as indicated by the lack of a mutation in the *psbA* gene and rapid accumulation of polar metabolites with the same retention time as a synthetic GSH-atrazine standard (Ma et al., 2013). Therefore, we hypothesize that the rapid formation of this metabolite results from increased GST activity in the resistant compared to sensitive populations, and that this increased activity is a result of higher expression of GST(s) or the presence of novel

GSTs with greater affinity towards atrazine. Therefore, the objectives of this study were to determine if differences in GST activity and expression of specific GST(s) existed between the atrazine-resistant and sensitive populations, as well as obtain protein and cDNA sequences from GSTs that confer resistance to atrazine in waterhemp.

Materials and Methods

Plant populations

Waterhemp populations used in these experiments were from McLean County (atrazine-resistant [MCR]), Adams County (atrazine-resistant [ACR]), and Wayne County (atrazine-sensitive [WCS]) Illinois, as described in (Hausman et al., 2011, Ma et al., 2013). Plants were germinated from seeds planted in 12 x 12-cm containers with a commercial potting medium (Sun Gro Horticulture) in a growth chamber. Waterhemp seedlings (2-cm) were then transferred to 80-cm³ containers in the greenhouse containing the same soil medium. When seedlings reached 4-cm, plants were transferred 950-cm³ pots with a 3:1:1:1 mixture of potting mix:soil:peat:sand, and a slow-release fertilizer was then added. Plants were then harvested (10-12 cm in height) for subsequent protein and enzyme extraction. Growth chamber and greenhouse conditions were maintained at 28°C/22°C day/night with 16/8-h photoperiod. Growth chamber light was provided by incandescent and fluorescent bulbs delivering 550 $\mu\text{mol m}^{-2}\text{s}^{-1}$ photon flux at plant canopy level. Natural light in the greenhouse was supplemented with mercury halide lamps, with a minimum output of 500 $\mu\text{mol m}^{-2}\text{s}^{-1}$ photon flux at the plant canopy level.

Crude protein extraction

Plants from each population were harvested as described above and stems were removed due to their interference with mechanical disruption. All plant leaves (including

the attached petiole) were frozen in liquid nitrogen and ground to a powder.

Polyvinylpolypyrrolidone (7.5% w/v) was added along with 3 mL extraction buffer (100 mM Tris-HCl (pH 7.0), 1 mM DTT, 1 mM EDTA) per gram fresh weight. Samples were then filtered through 3 layers of cheesecloth and centrifuged at 3,000 $\times$ g for 30 min at 5°C to obtain a clarified lysate.

Ammonium sulfate purification

Clarified lysates were used to prepare 40-80% ammonium sulfate (AMS) fractions based on protein solubility in water at 5°C. Initial experiments determined that more than 90% of GST activity (with atrazine as substrate) precipitated within this range, which is similar to previously published reports of plant GST purification (Riechers et al., 1997; Smith et al., 2004; Edwards and Dixon, 2005). After slowly adding AMS at each fractionation step, samples were stirred for 20 min at 5°C and then centrifuged at 3,000 $\times$ g for 30 min, at 5°C. Pelleted material from the 40% cut was discarded, as well as soluble fraction following the 80% cut. The remaining pellet from the 80% cut was then desalted in protein storage buffer (100 mM Tris (pH 7.0), 1mM DTT, 1mM EDTA) and stored at -80°C until further use.

GSH affinity purification and SDS-PAGE analysis

Desalted and concentrated protein samples from each population obtained after AMS precipitation were then further purified by immobilized GST affinity chromatography (General Electric, GSTrap FF column). A modified protocol, based on

the manufacturer's protocol, was used as described below. The column was equilibrated with 5 column volumes of binding buffer (140 mM NaCl, 10 mM NaHPO₄, 1.8 mM KHPO₄, and 2.7 mM KCl, pH 7.4). Protein samples (0.1-3.0 mg) were then loaded on the column and washed with 10 column volumes of binding buffer at a flow rate of 5 mL/min. Bound protein was then eluted with 10 column volumes of elution buffer (50 mM Tris-HCl (pH8.0), 10 mM GSH, 1 mM DTT) at a flow rate of 5 mL/min. Samples were then immediately desalted into protein storage buffer and analyzed by SDS-PAGE. Protein concentration was measured using a Nanodrop spectrophotometer using bovine serum albumin (BSA) as a standard. SDS-PAGE (12% separating gels) was conducted using the manufacturer's recommended (Bio-Rad) to examine purity of each protein fraction. Gels were silver- or Coomassie-stained using the manufacturer's recommended protocol (Bio-Rad).

GST-atrazine activity assay

To measure GST activity towards atrazine, a modified version of a previously published protocol was used (Gronwald and Plaisance, 1998). Reactions (300 µL) contained 50 µM ¹⁴C-atrazine (2 mCi/mmol specific activity), 2 mM GSH, 100 mM sodium citrate (pH 6.5), 1 mM 2-mercaptoethanol, 0.2 mg/mL BSA and protein fraction with a final concentration ranging from 0.05-1 mg/mL. A negative control was run for each replicate by replacing the protein fraction with extraction buffer. After incubation for 5 min at 30°C, reactions were initiated by the addition of ¹⁴C-atrazine and then incubated at 30°C for 30 min. Initial experiments with these assay conditions determined that product formation was linear until this time point. Reactions were

terminated by the addition of 50 μ L acetic acid and then partitioned with 900 μ L methylene chloride to separate unreacted atrazine with the atrazine-GSH conjugate. The amount of radioactivity in a 200 μ L aliquot of the aqueous phase was counted by liquid scintillation spectrometry (LSS). Non-enzymatic conjugation rates were determined by averaging radioactivity found in the aqueous phase of the negative control reactions (*i.e.*, no protein added) and was then subtracted from the total activity measured in the experimental reactions. Specific activity was then determined based on the protein concentration and the non-enzymatically corrected radioactivity quantified in the aqueous phase by LSS. Units of GST activity are reported as pmol of GSH-conjugated atrazine per minute per mg protein.

LC-MS/MS and bioinformatic analysis of peptide sequences

Coomassie-stained protein bands, representing proteins from each population that eluted from GSH affinity columns ($\sim$ 100ng), were manually excised from one-dimensional SDS-PAGE gels. Protein fractions were analyzed using nanoLC-MS/MS by the Stanford University Mass Spectrometry facility (333 Campus Drive, Stanford, CA). Samples were digested with trypsin and the resulting peptide fragments were purified and concentrated before loading onto a nanoLC C₁₈ analytical column. Charged peptides were then selected by abundance and sequenced by LC-MS/MS. The *Arabidopsis* reference genome was then searched with peptide sequences identified from each waterhemp population, and the resulting data was uploaded to Scaffold (Proteome Software, Inc; <http://www.proteomesoftware.com/products/scaffold>).

Database searches

All MS/MS samples were analyzed using Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version 1.0). Sequest was set up to search the NCBI_*Arabidopsis thaliana* database (unknown version, 222041 entries) assuming the digestion enzyme strict trypsin. Sequest was searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 20 PPM. Propionamide of cysteine was specified in Sequest as a fixed modification. Oxidation of methionine was specified in Sequest as a variable modification.

Criteria for protein identification

Scaffold (version Scaffold_4.3.4, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Sequest identifications required a minimum of deltaCn scores greater than 0.10 and XCorr scores greater than, 2.3, 2.4 and 2.4 for, doubly, triply and quadruply charged peptides. Protein identifications were accepted if they contained at least one identified peptide (Table A1). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were then grouped into clusters (Table A2). Proteins were annotated with gene ontology terms from NCBI (Ashburner et al., 2000).

Protein target selection

Arabidopsis proteins with matching peptide sequences were sorted based on protein identification probability. Protein matches greater than 90% (based on amino acid identity) were then screened by excluding proteins outside the target range for typical plant GST subunits (23-32 kDa) and proteins without an annotated function in the database. After screening, only proteins annotated as GSTs remained. All identical accession numbers for identified GSTs were obtained through the National Center for Biotechnology Information's Basic Local Alignment Search tool. These accession numbers were then used to search the annotations of the partially-sequenced transcriptome from the ACR waterhemp population (Lee et al., 2009) (individual reads available at <http://www.ncbi.nlm.nih.gov/sra/?term=amaranthus+tuberculatus>). Gene-specific primers (Table 2) were then designed for each matching waterhemp contig (Table A3).

Total RNA extraction and cDNA construction

Total RNA was extracted from frozen waterhemp leaves (as described earlier in Methods) using previously described methods (Riechers et al., 2003). Precipitated RNAs were washed twice with 75% ethanol and resuspended in diethylpyrocarbonate-treated water. Total RNA concentrations were determined using a NanoDrop spectrometer and RNA quality was confirmed by analysis with 1.2% agarose-formaldehyde denaturing gels as described previously (Riechers et al., 2003). First-strand cDNA synthesis was performed using the Maxima H-Minus cDNA synthesis kit

(Thermo-Scientific) following the manufacturer's protocol using 500 ng total RNA. The following parameters were then used for RT-PCR with 1 μ L of the cDNA reaction product: an initial denaturing step at 95°C for 4.5 min, then 30 amplification cycles consisting of 95°C for 50 s, 56°C for 55 s, and 72°C for 1 min. PCR was then concluded with a final extension step of 72°C for 8.5 min. PCR products were visualized with a 1% agarose gel and ethidium bromide staining.

qRT-PCR analysis of gene expression

Gene-specific primers were designed to amplify each waterhemp GST identified, and a beta-tubulin reference gene was selected from several annotated tubulin contigs (Table A3) from ACR (Lee et al., 2009) and determined as a suitable constitutive control gene for GST expression analyses. Reactions were set up in 20 μ L volumes following the manufacturer's protocol (Syber® Green RNA-to C_T[™] 1-Step Kit (Applied Biosystems)). The qRT-PCR protocol was as follows: 48°C 30min, 95°C 10min, 40 cycles of 95°C 15 s, 60°C 1 min, and a melting curve of 95°C 15 s, 60°C 15 s, 95°C 15 s. Dissociation curves for every reaction were analyzed to ensure only one replicon was amplified. Gene expression was calculated relative to WCS transcript levels in each sample using the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001).

Statistical Analysis

All experiments were independently conducted a minimum of three times. Enzymatic activity data were combined from two representative experiments with three

technical replications. qRT-PCR data are the result of three biological samples with three technical replications for each sample. ANOVA followed by Tukey's multiple comparison test ($P=0.05$) was conducted to determine significant differences with each experiment.

Results

GST activity towards atrazine in waterhemp populations

To determine potential differences in GST activity towards atrazine among the three waterhemp populations, specific activity (pmol/min/mg protein) was measured in crude plant extracts, AMS-precipitated fractions, and GSH affinity-purified fractions using radiolabeled atrazine (Table 1). There was no significant difference in total or specific GST activity in the resistant biotypes (ACR and MCR) when compared to WCS in crude protein extracts. However, specific GST activity towards atrazine among the AMS fractions showed higher activity in the two resistant biotypes when compared to the sensitive biotype, but no significant difference was observed between ACR and MCR. In the GSH affinity-purified fractions, specific GST activity was highest in MCR followed by ACR and WCS (Table 1). In addition, specific activity increased with each purification step, with crude fractions having the lowest activity and GSH affinity-purified fractions having the highest. The largest fold-increase in specific activity occurred during the GST affinity purification step with MCR, resulting in an approximate 9-fold purification relative to the crude extract (Table 1).

Significant differences in specific GST activity between the two atrazine-resistant and -sensitive populations were only found in partially-purified fractions (AMS and GSH affinity) (Table 1). This may be a result of the crude protein extracts containing many non-specific proteins, pigments, metabolites, or other cell debris, which could directly inhibit the GSH-atrazine conjugation reaction or by binding non-specifically with GST

enzymes. Since it was difficult to measure any significant differences in rates of GST-atrazine activities in crude plant extracts among waterhemp populations, a partial GST purification strategy was utilized. This purification scheme was based on previous methods used to measure the role of GSTs in herbicide tolerance in etiolated grain sorghum and wheat shoots (Riechers et al., 1997; Gronwald and Plaisance, 1998), in which proteins are AMS precipitated from crude protein extracts, then affinity purified. After AMS precipitation, the 40-80% precipitated fraction contained concentrated proteins that were readily soluble in buffer, with a significant removal of plant pigments, lipids, and cell debris clearly evident. Specific GST activities measured in these fractions displayed a trend that was expected based on prior atrazine-metabolism assays in these populations (Ma et al., 2013), with each resistant population having higher specific activity than WCS (Table 1). In addition, the specific GST activities in AMS fractions from ACR and MCR displayed yields that were higher than 100%, which is consistent with the presence of GST inhibitors present in crude plant extracts from these populations (Table 1). In order to further purify these fractions and attempt to isolate and identify unique or over-expressed GST isozymes in MCR and ACR, a GST affinity purification method was incorporated following AMS purification.

GSH affinity-purified protein fractions from each population showed a single, broad band at approximately 27 kDa when analyzed by SDS-PAGE (Figure 1), which is within the typical range (23-32 kDa) of plant GST subunits (Li et al., 2013) with few contaminating bands. However the GSH affinity-purified fractions (Figure 1) are likely a heterogeneous mixture of GSTs and other proteins with the ability to bind with GSH, such as GSH reductase. However, specific GST activity of these fractions followed the

same pattern as with the AMS fractions (Table 1), with each atrazine-resistant population having a higher specific activity than WCS. However, fold-purification levels for ACR and MCR (4.5 and 9.0, respectively, relative to crude extracts) were lower than those previously reported in cereal crops (Timmerman, 1989; Riechers et al., 1997; Zhang et al., 2004) for affinity-purified GSTs, which may be due to several reasons in our studies. Lower fold-purification levels in ACR and MCR may result from use of photosynthetic tissues instead of etiolated seedling or shoot tissues (Timmerman, 1989; Riechers et al., 1997; Zhang et al., 2004), loss of GST activity throughout sample processing following initial crude extract preparation through affinity chromatography, and/or from partially purifying GST proteins by selective binding of their active site via the GSH resin, which may have adversely affected GST catalytic activity when assayed later *in vitro*. However, the GST-atrazine activity found in these fractions can be attributed to the presence of several GST proteins, which as described below were confirmed to be primarily tau and phi-class GSTs through LC-MS/MS.

LC-MS/MS

Excised SDS-PAGE gel slices containing the approx. 27 kDa protein bands from each population were frozen on dry ice and submitted for LC-MS/MS analysis and peptide identification (performed by the Stanford University Mass Spectrometry Laboratory). The peptide sequences obtained following trypsin digestion were aligned with the *Arabidopsis* reference genome (Figure A1). Out of the 89 *Arabidopsis* sequence matches, about 30 had reported molecular weights within the range of 23-31 kDa (the approximate size of the excised protein band). Based on protein identification

probability, a phi-class GST and a tau-class GST were the two most likely matches identified in each population (Figures A1, A2). The peptide sequence KVLDVYEARL matched an *Arabidopsis* protein annotated as glutathione S-transferase F3 (GI:330250548), a phi class GST, and was the most probable match. The peptide sequence RFWADYIDKK was the second most probable match; it was annotated as glutathione S-transferase tau 25 (GI:15220041). All identical accessions were identified using the NCBI database. Matching accessions were then identified from an annotated ACR partial transcriptome (Lee et al., 2009), resulting in five matched contigs (Table 2).

Amplification of *Amaranthus tuberculatus* GST contigs

RT-PCR was used to determine if the matched contigs found in the annotated ACR transcriptome were expressed in each waterhemp population. Amplification was tested in each waterhemp population using the primer pairs listed in Table 2. Primers were designed to amplify the largest section of the contig (200-800bp) possible and were named based on its putative function and class. *Amaranthus rudis* (*Ar*) was used to denote the waterhemp species name to prevent confusion with *Arabidopsis thaliana*, since *A. rudis* and *A. tuberculatus* are synonymous (Pratt, 2001). All of the contigs amplified via RT-PCR in each waterhemp population generated amplicons of the predicted size, except for *ArGSTU3*. Additionally, primer crosses between primers matching the same accession (*i.e.*, forward primer for *ArGSTF1* with reverse primer for *ArGSTF2*) were made to determine if individual waterhemp contigs originated from the same mRNA. Direct sequencing of the RT-PCR amplicon derived from the *ArGSTF2* contig in MCR confirmed its identity as a GST (data not shown). Additional primers

based on the SYBR Green protocol for qRT-PCR were specifically designed for amplifying portions of the remaining contigs (Table 3; Table A3) except for *ArGSTF2*, since the original primers (Table 2) met qRT-PCR standards described in the protocol. *ArGSTF1* was excluded from further qRT-PCR analysis due to the formation of multiple PCR products (data not shown).

qRT-PCR analysis of gene expression

To determine if the matched genes with the predicted product size and single amplicon (*ArGSTF2*, *ArGSTU1*, and *ArGSTU2*) differed in levels of constitutive expression, qRT-PCR was performed using total RNA isolated from each waterhemp population at the same growth stage described previously for the protein purification studies. GST transcript levels were determined relative to expression in WCS (Livak and Schmittgen, 2001). *ArGSTF2* transcript levels in ACR and MCR were about 1000-fold higher than in WCS, but there was no significant difference between *ArGSTF2* transcript levels in these two atrazine-resistant biotypes (Figure 2A). There were no significant differences in transcript levels among populations for *ArGSTU1* (Figure 2B). Transcript levels for *ArGSTU2* were significantly higher in MCR than in ACR, but not between MCR and WCS or WCS and ACR (Figure 2C). Although not statistically different, a trend was observed where the abundance of *ArGSTU1* and *ArGSTU2* transcripts in MCR was higher than in WCS (Figures 2B, 2C). However, it is important to note that absolute expression levels for these two tau-class GSTs were much lower (about 30 to 300-fold lower for *ArGSTU1* and *ArGSTU2*, respectively) when compared with *ArGSTF2* expression levels in ACR and MCR (Figure 2A).

Higher constitutive expression of *ArGSTF2* in each atrazine-resistant population relative to WCS (Figure 2A), suggest that this gene may contribute to higher levels of GST activity (Table 1) and more rapid metabolism of atrazine (Ma et al. 2013). This is also supported by its identification as the highest probability match from the protein samples that had greater specific activity towards atrazine in ACR and MCR (Table 1). Greater transcript abundance of this GST gene may contribute to elevated GST activity and higher levels of GSH-atrazine metabolites formed in these atrazine-resistant waterhemp populations compared with WCS (Ma et al., 2013), and ultimately the atrazine-resistant phenotype in ACR and MCR. However, there may be other non-target site-resistance (NTSR) mechanisms contributing to the atrazine-resistant waterhemp populations examined in our study. For example, although *ArGSTU1* and *ArGSTU2* showed no significant differences in expression between MCR and WCS (Figures 2B, 2C), the expression of some GSTs are inducible in response to different abiotic stress factors such as herbicides or safeners (Fuerst et al., 1993; DeRidder et al., 2002; Riechers et al., 2003, 2010; Yang et al., 2009; Li et al., 2013; Wang et al., 2013). Increased enzymatic or non-enzymatic conjugation of atrazine to GSH may also in part result from increased synthesis of cysteine and/or GSH due to increased activity of enzymes such as cysteine synthase and gamma-glutamylcysteine synthetase, respectively, which are involved in cellular GSH biosynthesis (Nemat Alla et al., 2008). Furthermore, increased levels of ROS-detoxifying enzymes may also contribute to atrazine resistance in waterhemp, since atrazine triggers a massive accumulation of ROS and lipid peroxidation in chloroplasts in the presence of light as a major component of its mode of action (Hess, 2000).

Conclusions

Previous studies indicated that two different atrazine-resistant populations of waterhemp from Illinois result from non-target site resistance (NTSR) mechanism(s) (Ma et al., 2013). This hypothesis arose due to the lack of a mutation in *psbA* and the rapid accumulation of a polar metabolite with the same retention time via HPLC as a synthetic GSH-atrazine standard, which suggests that GSTs may be involved in atrazine resistance in ACR and MCR (Ma et al., 2013). Elevated basal GST activity with atrazine as a substrate was investigated as a potential NTSR mechanism in ACR and MCR in my current research.

Preliminary studies of shoots and roots, using both atrazine and synthetic GST substrates such as 1-chloro-2,4-dinitrobenzene (CDNB), narrowed our investigation of elevated GST activity to the leaves and petiole of waterhemp plants (data not shown). Both atrazine-resistant populations (ACR and MCR) had higher GST specific activity towards atrazine than WCS in AMS and GSH affinity-purified fractions. Proteins identified by LC-MS/MS from the GSH affinity-purified fractions were confirmed as primarily consisting of phi and tau-class GSTs by sequence similarity with *Arabidopsis* GSTs, and some of these GSTs proteins were successfully matched with cDNA contigs from a partial ACR transcriptome (Lee et al., 2009). One contig, identified as a phi-class GST (*ArGSTF2*), displayed higher constitutive expression in both atrazine-resistant waterhemp populations. The findings presented herein support the conclusion that herbicide-detoxifying GSTs are associated with atrazine resistance in MCR and ACR.

The development and characterization of F₂ populations in which atrazine resistance segregates, derived from a cross between WCS and either ACR or MCR, is

an important step in continuing this mechanistic research. Examining expression of *ArGSTF2* in these segregating populations and determining if a correlation exists with atrazine resistance will assist in determining if elevated expression of *ArGSTF2* is indeed a major genetic factor in conferring atrazine resistance in waterhemp. If a positive correlation between *ArGSTF2* expression and atrazine resistance is identified in additional *Amaranthus* populations or segregating F₂ lines, then this knowledge would provide a powerful tool to further understand the genetics and underlying molecular mechanisms of atrazine resistance. This approach would be similar to how segregating F₂ populations were used to discover a point mutation in the acetyl-CoA carboxylase gene from *Avena fatua* that resulted in resistance to diclofop and sethoxydim (Christoffers et al., 2006).

Preliminary experiments with an MCR x WCS F₁ hybrid and F₂ segregating lines suggest that atrazine resistance may be a single-gene trait, based on segregation analysis of the atrazine-resistant phenotype in F₂ lines (Huffman et al., unpublished). Preliminary qRT-PCR data support a role of *ArGSTF2* as the possible single genetic resistance factor, since its expression correlates with atrazine resistance in both resistant populations and resistant F₂ lines (Figure A1). Moreover, sensitive F₂ plants exhibit transcript levels similar to WCS, and F₁ plants displayed intermediate transcript levels. However, whether *ArGSTF2* is the predominant GST that confers atrazine resistance, or if other GSTs and/or NTSR mechanisms also contribute to atrazine resistance in MCR and ACR, is still unknown.

Global transcriptional analyses of these segregating F₂ populations, in which the whole transcriptome could be sequenced and number of GST genes quantified (Chi et

al., 2011), may enable further confirmation of the specific GST gene(s) correlating with resistance. Analysis of untreated and atrazine-treated waterhemp plants (sensitive and atrazine-resistant) may result in the identification of both constitutive and induced factors involved. If RNA-seq analysis indicated that *ArGSTF2* (or other metabolic enzymes) are the major genetic factor(s) correlating with resistance, then further biochemical and molecular characterization could be performed. This would require obtaining the entire open reading frame for expression and biochemical analyses of the recombinant GST enzyme, and allow for genomic cloning of waterhemp GSTs and subsequent bioinformatic comparisons between other plant GSTs conferring resistance to herbicides (Pei et al., 2014).

A similar approach was used in discovering a key mutation associated with metabolic resistance to dichlorodiphenyltrichloroethane (DDT) in mosquitos (*Anopheles funestus*). A genome-wide transcriptional analysis was conducted in which the most up-regulated gene was identified as a GST. This GST was then confirmed to confer resistance to DDT through transgenic expression analyses in sensitive *Drosophila*, and was determined to result from a point mutation in the wild-type GST gene in which leucine was substituted for phenylalanine (Riveron et al., 2014).

In weedy species without entire genomic sequences available, RNA-seq is being utilized for the characterization of evolved herbicide resistance mechanisms. Like waterhemp, annual ryegrass (*Lolium rigidum*) populations from Australia are known for their unique ability to evolve resistance to a wide variety of herbicides. NTSR mechanisms, likely involving P450s, towards acetyl-CoA carboxylase inhibitors was shown to arise in as few as three generations when grown under herbicide stress (Neve

and Powles, 2005; Yu et al., 2013; Gaines et al., 2014). Even though *Lolium rigidum* is a non-model plant species without a fully sequenced genome, the use of pyrosequencing to create a reference transcriptome allowed for subsequent identification and quantitative analysis of transcripts by Illumina sequencing (Gaines et al., 2014). Comparison of plants with evolved resistance to acetyl-CoA carboxylase inhibitors to their sensitive progenitors resulted in the identification of four genes with increased constitutive expression levels: two P450s, a nitronate monooxygenase, and a glucosyltransferase (Gaines et al., 2014). Additionally, one phi and two tau-class GST-annotated contigs were expressed higher in response to the synthetic auxin herbicide 2,4-D (Gaines et al., 2014).

In continuing this research, a method similar to one conducted in *Lolium* could be adopted using the partially-sequenced transcriptome from ACR (Lee et al., 2009) along with the sequenced grain amaranth (*Amaranthus hypochondriacus*) genome and transcriptome (Sunil et al., 2014) as references for RNA-seq. So far, the higher GST activity with atrazine quantified in ACR and MCR extracts can only be associated with higher constitutive expression of *ArGSTF2*. This observed difference could be due to a mutation or insertion resulting in a more active promoter. However, it is known that plant GSTs are found in gene clusters within the genome (Frova, 2006). An RNA-seq approach would allow for the identification of all genes with higher constitutive expression and also allow for the identification of potentially inducible genes in response to atrazine-induced oxidative stress. This would allow for a more comprehensive understanding of NTSR mechanism(s) to atrazine in waterhemp and other *Amaranthus* species.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. *Nature Genet* 25: 25-29.

Bowles ID, Lim EJ, Poppenberger B (2005) Glycosyltransferases: managers of small molecules. *Curr Opin Plant Biol* 8: 254-263.

Buchanan BB (2000) *Biochemistry and molecular biology of plants*. Rockville: American Society of Plant Physiologists. 1367 p

Chi Y, Cheng Y, Vanitha J, Kumar N, Ramamoorthy R, Ramachandran S, Jiang SY (2011) Expansion mechanisms and functional divergence of the glutathione S-transferase family in sorghum and other higher plants. *DNA Res* 18: 1-16.

Christoffers MJ, Berg ML, Messersmith CG (2002) An isoleucine to leucine mutation in acetyl-CoA carboxylase confers herbicide resistance in wild oat. *Genome* 45: 1049-1056.

Cummins I, Dixon DP, Freitag-Pohl S, Skipsey M, Edwards R (2011) Multiple roles for plant glutathione transferases in xenobiotic detoxification. *Drug Metab Rev* 43: 266-280.

Cummins I, Wortley DJ, Sabbadin F, He Z, Coxon CR, Straker HE, Edwards R (2013) Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. *Proc Natl Acad Sci USA* 110: 5812-5817.

DeRidder BP, Goldsbrough PB (2006) Organ-specific expression of glutathione S-transferases and the efficacy of herbicide safeners in *Arabidopsis*. *Plant Physiol* 140: 167-175.

DeRidder BP, Dixon DP, Beussman DJ, Edwards R, Goldsbrough PB (2002) Induction of glutathione S-transferases in *Arabidopsis* by herbicide safeners. *Plant Physiol* 130: 1497-1505.

Dixon DP, Skipsey M, Edwards R (2010) Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry* 71: 338-350.

Edwards R, Dixon DP, Walbot V (2000) Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci* 5: 193-198.

Edwards R, Dixon DP (2005) Plant glutathione transferases. *Methods Enzymol* 401: 169-186.

Frova C (2003) The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. *Physiol Plant* 119: 469-479.

Frova C (2006) Glutathione transferases in the genomics era: new insights and perspectives. *Biomol Eng* 23: 149-169.

Fuerst EP, Irzyk GP, Miller KD (1993) Partial characterization of glutathione S-transferase isozymes induced by the herbicide safener benoxacor in maize. *Plant Physiol* 102: 795–802.

Gaillard CD, Dufaud A, Tommasini R, Kreuz K, Amrhein N, Martinoia E (1994) A herbicide antidote (safener) induces the activity of both the herbicide detoxifying enzyme and of a vacuolar transporter for the detoxified herbicide. *FEBS Lett* 352: 219-221.

Gaines TA, Lorentz L, Figge A, Herrmann J, Maiwald F, Ott MC, Han H, Busi R, Yu Q, Powles SB, Beffa R (2014) RNA-Seq transcriptome analysis to identify genes involved in metabolism-based diclofop resistance in *Lolium rigidum*. *Plant J* 78: 865-876.

Gill SS, Anjum NA, Hasanuzzaman M, Gill R, Trivedi DK, Ahmad I, Pereira E, Tuteja N (2013) Glutathione and glutathione reductase: a boon in disguise for plant abiotic stress defense operations. *Plant Physiol Biochem* 70: 204-212.

Gingrich JC, Buzby JS, Stirewalt VL, Bryant DA (1988) Genetic analysis of two new mutations resulting in herbicide resistance in the cyanobacterium *Synechococcus* sp. PCC 7002. *Photosyn Res* 16: 83-99.

Girardini J, Amirante A, Zemzoumi K, Serra E (2002) Characterization of an omega-class glutathione S-transferase from *Schistosoma mansoni* with glutaredoxin-like dehydroascorbate reductase and thiol transferase activities. Eur J Biochem 269: 5512-5521.

Gray JA, Balke NE, Stoltenberg DE (1996) Increased glutathione conjugation of atrazine confers resistance in a Wisconsin velvetleaf (*Abutilon theophrasti*) biotype. Pesticide Biochem Physiol 55: 157-171.

Gronwald JW, Plaisance L (1998) Isolation and characterization of glutathione S-transferase isoenzymes from sorghum. Plant Physiol 117: 877-892.

Hager AG, Wax LM, Stoller EW, Bollero GA (2002) Common waterhemp (*Amaranthus rudis*) interference in soybean. Weed Sci 50: 607-610

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. Pest Manag Sci 67: 258-261.

Heap I (2014) The international survey of herbicide resistant weeds. <http://www.weedscience.org> (accessed April 29, 2014).

Hess F (2000). Light-dependent herbicides: an overview. Weed Sci 48: 160-170.

Hirschberg JM, McIntosh L (1983) Molecular basis of herbicide resistance in *Amaranthus hybridus*. Science 222: 1346-1349.

Holt DC, Lay VJ, Clarke ED, Dinsmore A, Jepson I, Bright SWJ, Greenland AJ (1995) Characterization of the safener-induced glutathione S-transferase isoform II from maize. Planta 196: 295-302

Holt JS, Powles SB, Holtum JA (1993) Mechanisms and agronomic aspects of herbicide resistance. Annu Rev Plant Biol 44: 203-229.

Krutz LJ, Burke IC, Reddy KN, Zablotowicz RM, Price AJ (2009) Enhanced atrazine degradation: evidence for reduced residual weed control and a method for identifying adapted soils and predicting herbicide persistence. *Weed Sci* 57: 427-434.

LeBaron HM, McFarland JE, Burnside OC (2011) *The triazine herbicides*. San Diego: Elsevier. 600 p.

Lee RM, Thimmapuram J, Thinglum KA, Gong G, Hernandez AG, Wright CL, Kim RW, Mikel MA, Tranel PJ (2009) Sampling the waterhemp (*Amaranthus tuberculatus*) genome using pyrosequencing technology. *Weed Sci*. 57: 463-469.

Li G, Wu SG, Yu RX, Cang T, Chen LP, Zhao XP, Cai LM, Wu CX (2013) Identification and expression pattern of a glutathione S-transferase in *Echinochloa crus-galli*. *Weed Res* 53: 314-321.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta CT$ method. *Methods* 25: 402-408.

Ma R, Kaundun SS, Tranel PJ, Riggins CW, McGinness DL, Hager AG, Hawkes T, McIndoe E, Riechers DE (2013) Distinct detoxification mechanisms confer resistance to mesotrione and atrazine in a population of waterhemp. *Plant Physiol* 163: 363-377.

Nemat Alla MM, Badawi A-HM, Hassan NM, El-Bastawisy ZM, Badran EG (2008) Herbicide tolerance in maize is related to increased levels of glutathione and glutathione-associated enzymes. *Acta Physiol Plant* 30: 371-379.

Neve PA, Powles S (2005) Recurrent selection with reduced herbicide rates results in the rapid evolution of herbicide resistance in *Lolium rigidum*. *Theor Appl Genet* 110: 1154-1166.

Oettmeier W (1992) Herbicides of photosystem II. *Topics Photosyn* 11: 349-408.

Pei DD, Ding J, Duan Z, Li M, Feng Y, Li C (2014) Cloning and expression of a tomato glutathione S-transferase (GST) in *Escherichia coli*. *Afric J Biotechnol* 11: 6402-6408.

Pratt DB, Clark LG (2001) *Amaranthus rudis* and *A. tuberculatus*, one species or two? J Torrey Bot Soc 128: 282-296.

Riechers DE, Irzyk GP, Jone, SS, Fuerst EP (1997) Partial characterization of glutathione S-transferases from wheat (*Triticum* spp.) and purification of a safener-induced glutathione S-transferase from *Triticum tauschii*. Plant Physiol 114: 1461-1470.

Riechers DE, Zhang Q, Xu F, Vaughn KC (2003) Tissue-specific expression and localization of safener-induced glutathione S-transferase proteins in *Triticum tauschii*. Planta 217: 831-840.

Riechers DE, Kreuz K, Zhang Q (2010) Detoxification without intoxication: herbicide safeners activate plant defense gene expression. Plant Physiol 153: 3-13.

Riveron JM, Yunta C, Ibrahim SS, Djouaka R, Irving H, Menze BD, Ismail HM, Hemingway J, Ranson H, Armando A, Wondji CS (2014) A single mutation in the *GSTe2* gene allows tracking of metabolically based insecticide resistance in a major malaria vector. Genome Biol 15(2): R27.

Schultz T, van Eck L, Botha AM (2014) Phi-class glutathione-S-transferase is involved in *Dn1*-mediated resistance. Physiol Plant [DOI: 10.1111/ppl.12284].

Schulz BK, Kolukisaoglu HÜ (2006) Genomics of plant ABC transporters: the alphabet of photosynthetic life forms or just holes in membranes? FEBS Lett 580: 1010-1016.

Smith AP, DeRidder BP, Guo WJ, Seeley EH, Regnier FE, Goldsbrough PB (2004) Proteomic analysis of *Arabidopsis* glutathione S-transferases from benoxacor- and copper-treated seedlings. J Biol Chem 279: 26,098-26,104.

Soranzo N, Gorla MS, Mizzi L, De Toma G, Frova C (2004) Organisation and structural evolution of the rice glutathione S-transferase gene family. Mol Gen Genom 271: 511-521.

Steckel LE (2007) The dioecious *Amaranthus* spp.: here to stay. Weed Technol 21: 567-570.

Sunil M, Hariharan AK, Nayak S, Gupta S, Nambisan SR, Gupta RP, Panda B, Choudhary B, Srinivasan S (2014) The draft genome and transcriptome of *Amaranthus hypochondriacus*: a C4 dicot producing high-lysine edible pseudo-cereal. DNA Res [doi:10.1093/dnares/dsu021].

Thompson Jr, L (1972) Metabolism of chloro-S-triazine herbicides by *Panicum* and *Setaria*. Weed Sci 20: 584-587.

Timmerman KP (1989) Molecular characterization of corn glutathione S-transferase isozymes involved in herbicide detoxication. Physiol Plant 77: 465-471.

Wagner U, Edwards R, Dixon DP, Mauch F (2002) Probing the diversity of the *Arabidopsis* glutathione S-transferase gene family. Plant Mol Biol 49: 515-532.

Wang ZH, Huang S, Jia C, Liu J, Zhang J, Xu B, Jin Z (2013) Molecular cloning and expression of five glutathione S-transferase (GST) genes from banana (*Musa acuminata* L. AAA group, cv. Cavendish). Plant Cell Rep 32: 1373-1380.

Wolf AE, Dietz KJ, Schröder P (1996) Degradation of glutathione S-conjugates by a carboxypeptidase in the plant vacuole. FEBS Lett 384: 31-34.

Woodyard AJ, Bollero GA, Riechers DE (2009) Broadleaf weed management in corn utilizing synergistic postemergence herbicide combinations. Weed Technol 23: 513-518.

Yang X, Sun W, Liu JP, Liu YJ, Zeng QY (2009) Biochemical and physiological characterization of a tau class glutathione transferase from rice (*Oryza sativa*). Plant Physiol Biochem 47: 1061-1068.

Yu Q, Han H, Cawthray GR, Wang, SF, Powles SB (2013) Enhanced rates of herbicide metabolism in low herbicide-dose selected resistant *Lolium rigidum*. Plant Cell Environ 36: 818-827.

Yuan JS, Tranel PJ, Stewart Jr, CN (2007) Non-target-site herbicide resistance: a family business. Trends Plant Sci 12: 6-13.

Zhang Q, Riechers DE (2004) Proteomic characterization of herbicide safener-induced proteins in the coleoptile of *Triticum tauschii* seedlings. *Proteomics* 4: 2058-2071.

Tables and Figures

Table 1. Activity of GST proteins isolated from three different *Amaranthus tuberculatus* populations, using atrazine as a substrate.

Purification step	Total activity (pmol)			specific activity (pmol/min/mg)			Activity yield			Fold Purification		
	WCS	ACR	MCR	WCS	ACR	MCR	WCS	ACR	MCR	WCS	ACR	MCR
Crude	220.5	199.1	249.8	32.7±1.4	36.0±10.6	27.9±0.9	100.0%	100.0%	100.0%	1.0	1.0	1.0
AMS	87.0	353.8	263.9	33.5±5.5	86.9±7.3	101.9±8.6	39.5%	177.8%	105.6%	1.0	2.4	3.7
GSH	63.8	91.3	142.8	39.4±9.1	163.1±5.2	250.6±8.0	28.9%	45.9%	57.2%	1.2	4.5	9.0

Waterhemp plants were harvested at 10-12cm. A clarified lysate (Crude) was selectively precipitated using ammonium sulfate (AMS). The resulting pellet was then desalted into extraction buffer and then loaded on to an immobilized glutathione column. Bound protein (GSH) was eluted and the GST activity (pmol/min/mg protein) in each fraction was assayed with radiolabeled atrazine as a substrate. This data is the combination of two representative experiments. McLean County (atrazine-resistant [MCR]), Adams County (atrazine-resistant [ACR]), and Wayne County (atrazine-sensitive [WCS]) Illinois, as described in (Hausman et al. 2011; Ma et al. 2013).

Table 2. PCR primers for used for contig analysis.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
ArGSTF1	GCTAATGAGAAAGAGCTTCAATAC	ACTTACTACTGGATCAAATTGG
ArGSTF2	GCACCCAACGTGTATTAG	AGTAAGGGGTGTTCCCTTG
ArGSTU1	TGAGAAATAGTAGGAGAAATTTGGAC	CTCAACCCATACTTCAATCTACC
ArGSTU2	TCCTATCCCAAGATAAACTCAG	GGAGTATATTGATGAATATTGGAAAGG
ArGSTU3	CGCCAAGCATTGACATGAGG	AAACTCGCAATGGAGCTTGC

Primers based on matched contigs from the ACR transcriptome (Lee et al. 2009) to obtain the largest fragment possible.

Table 3. PCR primers for expression analysis.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
ArGSTF1	GCTAATGAGAAAGAGCTTCAATAC	ACTTACTACTGGATCAAATTGG
ArGSTF2	GCACCCAACGTGTATTAG	AGTAAGGGGTGTTCCCTTG
ArGSTU1	TCGAGTAGGGGAAACGCGGTA	CCAGACCCCTGCCCTCTAGCCGA
ArGSTU2	ACCCCAAGCAATGATTTTGGGACA	AGACACTAGAACAAAGTGTGGGAGA
βTubulin	AGATTTTTCGCCCGGATAAC	TCCCATTCCAGATCCCTGTTC

Primers were designed based on the parameters suggested in the Syber® Green RNA-to CT™ 1-Step Kit (Applied Biosystems).

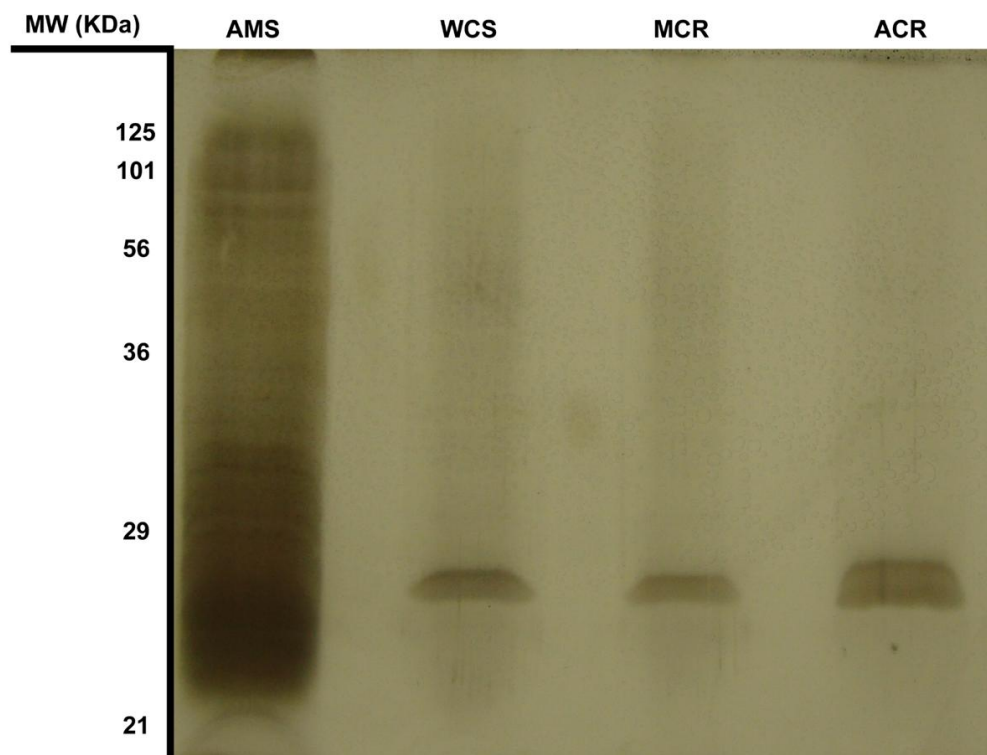


Figure 1. SDS-PAGE of isolated proteins from *Amaranthus tuberculatus* plants. Representative desalted AMS fractions from MCR followed by representative affinity purified eluents (GSH) from MCR, WCS, and ACR. Proteins were analyzed with 12% SDS-PAGE separating gels. Approximately 300ng AMS, and between 40-80ng GSH fractions were loaded per lane. Gels were silver stained as described in Methods. McLean County (atrazine-resistant [MCR]), Adams County (atrazine-resistant [ACR]), and Wayne County (atrazine-sensitive [WCS]) Illinois, as described in (Hausman et al. 2011; Ma et al. 2013).

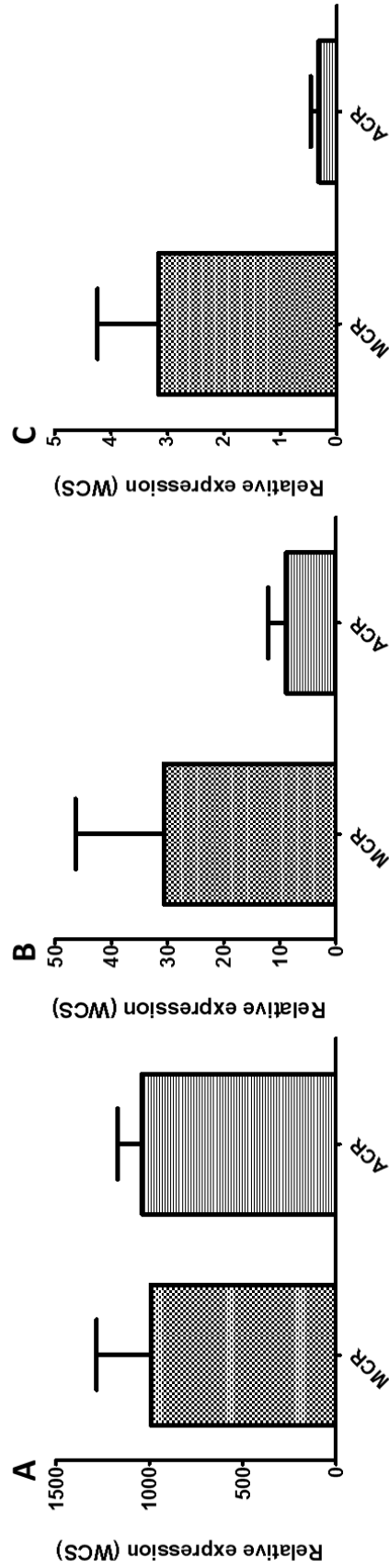


Figure 2. Expression of candidate GSTs in MCR and ACR relative to WCS. qRT-PCR assays were conducted for each target GST using primers from Table 3. Dissociation curves for every reaction were analyzed to ensure only one replicon was amplified. Expression of *ArGSTF2* (**A**), *ArGSTU1* (**B**), and *ArGSTU2* (**C**) in MCR and ACR was calculated relative to expression in WCS using the $2^{-\Delta\Delta CT}$ method (Livak 2001), with expression of beta-tubulin (Table A3) as a constitutive control gene. Data represent the mean of three biological replications with three technical replications each. Bars represent the standard error of the mean. McLean County (atrazine-resistant [MCR]), Adams County (atrazine-resistant [ACR]), and Wayne County (atrazine-sensitive [WCS]) Illinois, as described in (Hausman et al. 2011; Ma et al. 2013).

Appendix

#	Identified Proteins (89)	Accession Number	Molecular Weight	Protein Identification Probability			
				ACR	MCR	WCS	
1.1	RecName: Full=Glutathione S-transferase 16; AltName: F	gi 12230166 (+2)	24 kDa	93%	39%	98%	
1.2	RecName: Full=Glutathione S-transferase 6, chloroplastic	gi 148887379 (+3)	29 kDa	38%	32%	89%	
2	glutathione S-transferase TAU 23 [Arabidopsis thaliana]	gi 15218301 (+4)	26 kDa	100%	100%	100%	
3	unnamed protein product [Arabidopsis thaliana]	gi 10177920 (+6)	31 kDa	100%	100%	100%	
4	pyruvate, phosphate dikinase 1 [Arabidopsis thaliana]	gi 145333150 (+9)	104 kDa	100%	100%	100%	
5.1	phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana]	gi 332642071 (+1)	110 kDa	100%	100%	100%	
5.2	phosphoenolpyruvate carboxylase 2 [Arabidopsis thaliana]	gi 240254633 (+3)	110 kDa	0	100%	88%	
5.3	RecName: Full=Phosphoenolpyruvate carboxylase 1; Sho	gi 24636264 (+3)	110 kDa	16%	51%	32%	
6.1	RecName: Full=14-3-3-like protein GF14 lambda; AltName: F	gi 1345595	28 kDa	99%	98%	81%	
6.2	RecName: Full=14-3-3-like protein GF14 nu; AltName: F	gi 3023217	30 kDa	67%	100%	87%	
6.3	RecName: Full=14-3-3-like protein GF14 omega; AltName: g	gi 12230867	29 kDa	25%	54%	18%	
6.4	RecName: Full=14-3-3-like protein GF14 phi; AltName: F	gi 1702987	30 kDa	8%	30%	12%	
6.5	RecName: Full=14-3-3-like protein GF14 chi; AltName: F	gi 152112421	30 kDa	17%	63%	12%	
7.1	heat shock protein 70-4 [Arabidopsis thaliana]	gi 332641697 (+1)	71 kDa	20%	100%	0	
7.2	RecName: Full=Heat shock cognate 70 kDa protein 3; Sho	gi 18206367 (+2)	71 kDa	17%	100%	0	
7.3	RecName: Full=Heat shock cognate 70 kDa protein 2; Sho	gi 12644165	71 kDa	8%	68%	0	
7.4	RecName: Full=Heat shock cognate 70 kDa protein 1; Sho	gi 12643273	71 kDa	13%	96%	0	
7.5	RecName: Full=Luminal-binding protein 2; Short=AtBP2; g	gi 12643245 (+2)	74 kDa	12%	39%	0	
7.6	RecName: Full=Luminal-binding protein 1; Short=AtBP1; g	gi 18206379	74 kDa	8%	63%	0	
7.7	heat shock protein-70 cognate protein [Arabidopsis thali	gi 332195268	68 kDa	0	26%	0	
8.1	Transketolase [Arabidopsis thaliana]	gi 18411711 (+8)	80 kDa	100%	100%	100%	
8.2	Transketolase [Arabidopsis thaliana]	gi 30689983 (+1)	80 kDa	28%	43%	29%	
9.1	RecName: Full=Malate dehydrogenase 1, mitochondrial;	gi 11133715 (+2)	36 kDa	100%	99%	98%	
9.2	malate dehydrogenase 2 [Arabidopsis thaliana]	gi 145332399 (+4)	33 kDa	44%	39%	32%	
10.1	phosphoglycerate kinase [Arabidopsis thaliana]	gi 15223484 (+4)	50 kDa	100%	100%	100%	
10.2	phosphoglycerate kinase 1 [Arabidopsis thaliana]	gi 15230595 (+2)	50 kDa	77%	97%	99%	
11.1	fructose-bisphosphate aldolase, class I [Arabidopsis thali	gi 330252066 (+1)	43 kDa	100%	100%	100%	
11.2	fructose-bisphosphate aldolase, class I [Arabidopsis thali	gi 330250312 (+1)	42 kDa	97%	0	28%	
12	RecName: Full=ATP synthase subunit beta, chloroplastic; g	gi 6686269 (+1)	54 kDa	0	100%	0	
13	RecName: Full=Elongation factor 1-alpha; Short=EF-1- al	gi 119143 (+9)	50 kDa	100%	100%	100%	
14.1	ClpC [Arabidopsis thaliana]	gi 2921158	103 kDa	0	100%	58%	
14.2	Clp ATPase [Arabidopsis thaliana]	gi 186510816 (+4)	106 kDa	0	100%	0	
14.3	RecName: Full=ERD1 protein, chloroplasti; Flags: Precur	gi 1169544 (+2)	103 kDa	0	7%	0	
15.1	heat shock protein 81-2 [Arabidopsis thaliana]	gi 15241115 (+2)	80 kDa	0	100%	0	
15.2	RecName: Full=Heat shock protein 90-1; Short=AtHSP90.1	gi 26454635 (+1)	81 kDa	0	72%	0	
16.1	5-methyltetrahydropteroyltriglutamate--homocysteine r	gi 145334507 (+3)	84 kDa	64%	100%	87%	
16.2	methionine synthase 3 [Arabidopsis thaliana]	gi 145334531 (+2)	91 kDa	14%	48%	27%	
16.3	methionine synthase 2 [Arabidopsis thaliana]	gi 186509731 (+3)	85 kDa	13%	48%	24%	
17.1	glyceraldehyde-3-phosphate dehydrogenase (NADP+) (p	gi 186478427 (+4)	38 kDa	0	100%	0	
17.2	RecName: Full=Glyceraldehyde-3-phosphate dehydroge	gi 20455491 (+1)	48 kDa	0	100%	95%	
18	cystolic glyceraldehyde-3-phosphate dehydrogenase [Ar	gi 166706 (+1)	37 kDa	0	100%	90%	
19.1	RecName: Full=Malate dehydrogenase, cytoplasmic 2	gi 11133446 (+2)	36 kDa	61%	100%	66%	
19.2	RecName: Full=Malate dehydrogenase, cytoplasmic 1	gi 11133509 (+4)	36 kDa	43%	92%	64%	
20	RecName: Full=Proteasome subunit alpha type-5-A; AltN	gi 12229903 (+10)	26 kDa	84%	100%	100%	
21	putative ribulose-5-phosphate-3-epimerase [Arabidopsi	gi 15027997 (+7)	30 kDa	94%	91%	100%	
22	phosphoglycerate kinase, putative [Arabidopsis thaliana]	gi 21536853 (+1)	42 kDa	39%	100%	100%	
23	putative carboxymethylglutamate-kinase [Arabidopsis th	gi 110743666 (+5)	26 kDa	83%	100%	100%	
24	heat shock protein 70-2 [Arabidopsis thaliana]	gi 15240578 (+2)	77 kDa	0	100%	0	
25	NAD-dependent malic enzyme 1 [Arabidopsis thaliana]	gi 15225438 (+4)	70 kDa	0	100%	90%	
26.1	proteasome subunit beta type-7-B [Arabidopsis thaliana]	gi 30693626 (+5)	30 kDa	99%	92%	100%	
26.2	proteasome subunit beta type-7-A [Arabidopsis thaliana]	gi 332643797 (+4)	29 kDa	12%	0	6%	
27.1	RecName: Full=Proteasome subunit alpha type-6-B; AltN	gi 12229902 (+5)	27 kDa	42%	99%	75%	
27.2	RecName: Full=Proteasome subunit alpha type-6-A; AltN	gi 12643647 (+4)	27 kDa	0	0	0	
28	adenosylhomocysteinase 1 [Arabidopsis thaliana]	gi 145333043 (+4)	36 kDa	0	100%	90%	
29	thiol protease aleurain-like protein [Arabidopsis thaliana]	gi 332644500 (+3)	40 kDa	57%	100%	0	
30	Unknown protein [Arabidopsis thaliana]	gi 17065446 (+7)	34 kDa	0	100%	100%	
31.1	adenine phosphoribosyltransferase 1 [Arabidopsis thalia	gi 18396344 (+6)	26 kDa	0	75%	94%	
31.2	RecName: Full=Adenine phosphoribosyltransferase 2; Sh	gi 2499931 (+1)	21 kDa	0	18%	11%	
32	hypothetical protein F16M2.40 - Arabidopsis thaliana	gi 11277801 (+1)	32 kDa	97%	99%	99%	
33	RecName: Full=Eukaryotic initiation factor 4A-1; Short=el	gi 1170503 (+8)	47 kDa	0	100%	96%	
34	unnamed protein product [Arabidopsis thaliana]	gi 300648936 (+4)	60 kDa	94%	97%	96%	

Table A1. LC-MS/MS matches sorted by protein identification probability across samples from each waterhemp population (ACR, MCR, and WCS).

35	DING protein [Arabidopsis thaliana]	gi 53854257	6 kDa	100%	100%	0
36	2,3-bisphosphoglycerate-independent phosphoglycerate	gi 18391066 (+4)	61 kDa	0	100%	0
37	elongation factor G [Arabidopsis thaliana]	gi 62320893	45 kDa	92%	100%	0
38.1	S-adenosylmethionine synthase 4 [Arabidopsis thaliana]	gi 332642427 (+1)	43 kDa	0	100%	0
38.2	RecName: Full=S-adenosylmethionine synthase 2; Short:	gi 127045 (+2)	43 kDa	0	0	0
39	proteasome subunit beta type-6 [Arabidopsis thaliana]	gi 15235889 (+10)	25 kDa	0	100%	64%
40	RecName: Full=L-ascorbate peroxidase 2, cytosolic; AltN:	gi 118572828 (+10)	28 kDa	92%	86%	63%
41.1	RecName: Full=Chaperonin CPN60, mitochondrial; AltNa	gi 12644189 (+3)	61 kDa	0	97%	0
41.2	mitochondrial chaperonin HSP60 [Arabidopsis thaliana]	gi 22531044 (+6)	62 kDa	0	0	0
42	putative 20S proteasome beta subunit PBC2 [Arabidopsis	gi 12744997 (+10)	23 kDa	0	0	98%
43	peptide methionine sulfoxide reductase A3 [Arabidopsis	gi 15240796 (+2)	23 kDa	50%	99%	0
44	RecName: Full=DEAD-box ATP-dependent RNA helicase	gi 108861899 (+10)	48 kDa	0	91%	0
45.1	auxin-induced atb2-like protein [Arabidopsis thaliana]	gi 332195607	36 kDa	0	82%	98%
45.2	NAD(P)-linked oxidoreductase-like protein [Arabidopsis	gi 332195599	38 kDa	0	0	27%
45.3	NAD(P)-linked oxidoreductase-like protein [Arabidopsis	gi 30696459 (+3)	38 kDa	0	0	12%
46.1	60S ribosomal protein L8-3 [Arabidopsis thaliana]	gi 15234298 (+2)	28 kDa	0	57%	0
46.2	60S ribosomal protein L8-2 [Arabidopsis thaliana]	gi 186510913 (+2)	28 kDa	0	0	0
47	RecName: Full=Probable phosphoglucomutase, cytoplasi	gi 12585324 (+5)	63 kDa	0	100%	0
48	oxygen-evolving enhancer protein 3-2 [Arabidopsis thali	gi 15234637 (+3)	25 kDa	0	93%	0
49	peptide methionine sulfoxide reductase A2 [Arabidopsis	gi 15240795 (+4)	24 kDa	33%	100%	0
50	Glycosyl hydrolase family protein [Arabidopsis thaliana]	gi 22326920 (+3)	68 kDa	0	44%	0
51	60S ribosomal protein L13-1 [Arabidopsis thaliana]	gi 145332789 (+10)	24 kDa	0	0	0
52	chaperonin-60 alpha subunit [Arabidopsis thaliana]	gi 1223910 (+9)	62 kDa	0	100%	0
53	RecName: Full=Proteasome subunit alpha type-3; AltNan	gi 12644056 (+10)	27 kDa	0	0	0
54.1	RecName: Full=Ferredoxin-dependent glutamate syntha	gi 12643970 (+6)	180 kDa	0	0	0
54.2	ferredoxin-dependent glutamate synthase 2 [Arabidopsi	gi 15227306 (+3)	178 kDa	0	0	0
55.1	putative WD-40 repeat protein, MSI4 [Arabidopsis thaliar	gi 20148237 (+7)	56 kDa	0	0	0
55.2	histone-binding protein RBBP4 [Arabidopsis thaliana]	gi 332660266 (+2)	54 kDa	0	0	0

Table A1. Continued

Biological sample name	Protein name	Protein accession numbers	M.W. (Da)	Protein ID prop	Total spectrum count	Percentage of total spectra	Percentage sequence coverage
ACR	RecName: Full=Malate dehydrogenase 1, mitochondrial; AltName: Full=Heat shock protein 70-4 [Arabidopsis thaliana]	gi 11133715,gi 18404382,gi 332194789 gi 332641697,gi 75311168	35,805.20 71,103.20	99.70% 20.40%	3 0	0.02% 0.00%	10.30% 0.00%
ACR	Elongation factor 1-alpha [Arabidopsis thaliana]	gi 119143,gi 145323788,gi 186532608,gi 332009935,gi 33	49,503.40	99.90%	2	0.01%	4.68%
ACR	DING protein [Arabidopsis thaliana]	gi 53854257	5,554.20	99.90%	2	0.01%	49.10%
ACR	phosphoglycerate kinase [Arabidopsis thaliana]	gi 15223484,gi 186491325,gi 332195235,gi 332195236,gi	49,939.50	100.00%	4	0.03%	12.10%
ACR	fructose-bisphosphate aldolase, class I [Arabidopsis thaliana]	gi 330252066,gi 75313518	42,931.50	99.90%	2	0.01%	7.02%
ACR	D-ribulose-5-phosphate-3-epimerase [Arabidopsis thaliana]	gi 15027997,gi 15240250,gi 20259209,gi 30697525,gi 332	30,008.50	94.20%	1	0.01%	5.69%
ACR	At2g32520 [Arabidopsis thaliana]	gi 110743666,gi 15225693,gi 21554208,gi 30102612,gi 32	25,921.60	83.40%	0	0.00%	0.00%
ACR	proteasome subunit alpha type-5-B [Arabidopsis thaliana]	gi 12229903,gi 145325425,gi 15220961,gi 15231824,gi 18	25,947.30	83.90%	0	0.00%	0.00%
ACR	proteasome subunit beta type-7-B [Arabidopsis thaliana]	gi 30693626,gi 332007184,gi 332007185,gi 332007186,gi	29,617.60	99.20%	1	0.01%	5.11%
ACR	RecName: Full=14-3-3-like protein GF14lambda; AltName: Full=	gi 1345595	27,976.50	98.50%	2	0.01%	4.84%
ACR	transketolase-like protein [Arabidopsis thaliana]	gi 18411711,gi 20268682,gi 21553673,gi 22136900,gi 233	79,841.00	100.00%	1	0.01%	1.75%
ACR	unnamed protein product [Arabidopsis thaliana]	gi 300648936,gi 30698866,gi 332197184,gi 332197186,gi	59,512.40	94.20%	1	0.01%	1.85%
ACR	RecName: Full=Proteasome subunit alpha type-6-B; AltName: Full=	gi 12229902,gi 15224993,gi 300536463,gi 330250885,gi 3	27,350.30	41.70%	0	0.00%	0.00%
ACR	RecName: Full=Heat shock cognate 70kDa protein 3; Short=Hsc	gi 18206367,gi 332641246,gi 332641247	71,149.90	16.60%	0	0.00%	0.00%
ACR	phosphoglycerate kinase 1 [Arabidopsis thaliana]	gi 15230595,gi 332641725,gi 75273138	50,113.10	77.10%	3	0.02%	8.32%
ACR	RecName: Full=Malate dehydrogenase, cytoplasmic 2	gi 11133446,gi 15239843,gi 332007563	35,674.80	61.40%	2	0.01%	4.82%
ACR	AT5g47840/MCA23_18 [Arabidopsis thaliana]	gi 10177920,gi 15146328,gi 15238822,gi 227295722,gi 24	31,452.70	100.00%	5	0.03%	12.00%
ACR	peptide methionine sulfoxide reductase A3 [Arabidopsis thaliana]	gi 15240796,gi 332003779,gi 75264361	22,544.60	50.00%	0	0.00%	0.00%
ACR	peptide methionine sulfoxide reductase A2 [Arabidopsis thaliana]	gi 15240795,gi 332003778,gi 38454142,gi 46402480,gi 75	24,433.90	33.00%	0	0.00%	0.00%
ACR	elongation factor G [Arabidopsis thaliana]	gi 62320893	44,945.90	91.90%	0	0.00%	0.00%
ACR	pyruvate, phosphate dikinase 1 [Arabidopsis thaliana]	gi 145333150,gi 145333164,gi 186511835,gi 193806403,gi	105,140.10	100.00%	6	0.04%	4.50%
ACR	S-methyltetrahydropteroylglutamate--homocysteine methyl	gi 145334507,gi 332005103,gi 332005104,gi 8134566	84,358.40	63.60%	1	0.01%	1.31%
ACR	RecName: Full=14-3-3-like protein GF14nu; AltName: Full=Ger	gi 3023217	29,825.60	66.50%	0	0.00%	0.00%
ACR	hypothetical protein F16M2.40 - Arabidopsis thaliana	gi 11277801,gi 7523401	32,131.10	97.20%	1	0.01%	3.46%
ACR	phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana]	gi 332642071,gi 73917651	110,164.10	99.60%	2	0.01%	2.38%
ACR	RecName: Full=Ascorbate peroxidase 2, cytosolic; AltName: Full=	gi 118572828,gi 257352621,gi 257632823,gi 257643578,gi	27,947.00	91.80%	0	0.00%	0.00%
ACR	thiol protease aleurain-like protein [Arabidopsis thaliana]	gi 332644500,gi 332644501,gi 73622182,gi 79314271	39,541.60	57.40%	0	0.00%	0.00%
★	glutathione S-transferase TAU 23 [Arabidopsis thaliana]	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 944	25,682.10	99.80%	3	0.02%	4.09%
ACR	phosphoglycerate kinase, putative [Arabidopsis thaliana]	gi 21536853,gi 298541591	42,148.90	38.70%	0	0.00%	0.00%
MCR	RecName: Full=Malate dehydrogenase 1, mitochondrial; AltName: Full=	gi 11133715,gi 18404382,gi 332194789	35,805.20	99.40%	5	0.03%	10.60%
MCR	60S ribosomal protein L8-3 [Arabidopsis thaliana]	gi 15234298,gi 332661223,gi 75221276	27,948.20	57.40%	0	0.00%	0.00%
MCR	heat shock protein 70-4 [Arabidopsis thaliana]	gi 332641697,gi 75311168	71,103.20	99.90%	11	0.07%	17.70%
MCR	Elongation factor 1-alpha [Arabidopsis thaliana]	gi 119143,gi 145323788,gi 186532608,gi 332009935,gi 33	49,503.40	100.00%	3	0.02%	6.46%
MCR	DING protein [Arabidopsis thaliana]	gi 53854257	5,554.20	100.00%	2	0.01%	54.70%
MCR	phosphoglycerate kinase [Arabidopsis thaliana]	gi 15223484,gi 186491325,gi 332195235,gi 332195236,gi	49,939.50	100.00%	3	0.02%	8.58%
MCR	adenosylhomocysteinease 1 [Arabidopsis thaliana]	gi 145333043,gi 15236376,gi 332657945,gi 332657948,gi	53,379.20	100.00%	4	0.02%	8.45%
MCR	fructose-bisphosphate aldolase, class I [Arabidopsis thaliana]	gi 330252066,gi 75313518	42,931.50	100.00%	3	0.02%	10.80%
MCR	D-ribulose-5-phosphate-3-epimerase [Arabidopsis thaliana]	gi 15027997,gi 15240250,gi 20259209,gi 30697525,gi 332	30,008.50	90.70%	0	0.00%	0.00%
MCR	At2g32520 [Arabidopsis thaliana]	gi 110743666,gi 15225693,gi 21554208,gi 30102612,gi 32	25,921.60	99.80%	1	0.01%	5.02%
MCR	putative malate oxidoreductase (malic enzyme) [Arabidopsis thaliana]	gi 15225438,gi 20268709,gi 24030498,gi 330251148,gi 45	69,657.30	99.70%	2	0.01%	2.25%
MCR	Unknown protein [Arabidopsis thaliana]	gi 17065446,gi 18399513,gi 20148535,gi 21592994,gi 330	33,797.20	99.60%	1	0.01%	2.67%
MCR	phosphoglycerate kinase 1 [Arabidopsis thaliana]	gi 15230595,gi 332641725,gi 75273138	50,113.10	96.50%	2	0.01%	4.78%
MCR	proteasome subunit alpha type-5-B [Arabidopsis thaliana]	gi 12229903,gi 145325425,gi 15220961,gi 15231824,gi 18	25,947.30	100.00%	2	0.01%	10.50%
MCR	Glycosyl hydrolase family protein [Arabidopsis thaliana]	gi 22326920,gi 30688067,gi 332005527,gi 332005528	67,927.80	44.20%	0	0.00%	0.00%
MCR	glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosph	gi 186478427,gi 332190823,gi 332190825,gi 332190826,gi	37,667.30	100.00%	5	0.03%	14.30%
MCR	Ctp ATPase [Arabidopsis thaliana]	gi 186510816,gi 296529804,gi 332644945,gi 332644946,gi	105,774.70	100.00%	5	0.03%	5.67%
MCR	2,3-bisphosphoglycerate-independent phosphoglycerate mut	gi 18391066,gi 19548023,gi 2160168,gi 30216342,gi 3321	62,669.60	100.00%	2	0.01%	3.05%
MCR	RecName: Full=Eukaryotic initiation factor 4A-2; Short=elf-4A-	gi 1170503,gi 1170505,gi 332194952,gi 332194953,gi 332	46,762.90	100.00%	2	0.01%	5.83%
MCR	RecName: Full=Proteasome subunit alpha type-6-B; AltName: Full=	gi 12229902,gi 15224993,gi 300536463,gi 330250885,gi 3	27,350.30	98.90%	1	0.01%	3.66%
MCR	RecName: Full=Glyceraldehyde-3-phosphate dehydrogenase f	gi 20455491,gi 332193812	47,659.30	100.00%	4	0.02%	11.90%
MCR	RecName: Full=ATP synthase subunit beta, chloroplastic; AltName: Full=	gi 6686269,gi 7525040	53,934.80	100.00%	15	0.09%	27.50%
MCR	RecName: Full=14-3-3-like protein GF14lambda; AltName: Full=	gi 1345595	27,976.50	98.00%	1	0.01%	4.03%
MCR	At4g05180/C17L7_100 [Arabidopsis thaliana]	gi 15234637,gi 18206249,gi 23308489,gi 332657088	24,643.90	93.00%	1	0.01%	4.35%
MCR	transketolase-like protein [Arabidopsis thaliana]	gi 18411711,gi 20268682,gi 21553673,gi 22136900,gi 233	79,841.00	100.00%	4	0.02%	9.45%
MCR	unnamed protein product [Arabidopsis thaliana]	gi 300648936,gi 30698866,gi 332197184,gi 332197186,gi	59,512.40	96.70%	1	0.01%	1.85%
MCR	RecName: Full=Heat shock cognate 70kDa protein 3; Short=Hsc	gi 18206367,gi 332641246,gi 332641247	71,149.90	100.00%	15	0.09%	15.90%
MCR	thiol protease aleurain-like protein [Arabidopsis thaliana]	gi 332644500,gi 332644501,gi 73622182,gi 79314271	39,541.60	100.00%	2	0.01%	7.26%
MCR	AT5g47840/MCA23_18 [Arabidopsis thaliana]	gi 10177920,gi 15146328,gi 15238822,gi 227295722,gi 24	31,452.70	100.00%	4	0.02%	12.00%
MCR	proteasome subunit beta type-7-B [Arabidopsis thaliana]	gi 30693626,gi 332007184,gi 332007185,gi 332007186,gi	29,617.60	92.00%	1	0.01%	3.28%
MCR	phosphoenolpyruvate carboxylase 2 [Arabidopsis thaliana]	gi 240254633,gi 330255051,gi 330255052,gi 73917650	109,757.20	99.60%	6	0.04%	4.67%
MCR	DEAD-box ATP-dependent RNA helicase 56 [Arabidopsis thaliana]	gi 108861899,gi 186522071,gi 186522075,gi 238481238,gi	52,860.10	91.40%	1	0.01%	2.34%
MCR	peptide methionine sulfoxide reductase A3 [Arabidopsis thaliana]	gi 15240796,gi 332003779,gi 75264361	22,544.60	99.00%	2	0.01%	10.90%
MCR	peptide methionine sulfoxide reductase A2 [Arabidopsis thaliana]	gi 15240795,gi 332003778,gi 38454142,gi 46402480,gi 75	24,433.90	99.80%	2	0.01%	10.10%
MCR	RecName: Full=Probable phosphoglucosyltransferase 2, cytoplasmic 2; g	gi 12585324,gi 238479031,gi 238479033,gi 332196984,gi	63,483.50	100.00%	2	0.01%	3.59%
MCR	S-adenosylmethionine synthase 4 [Arabidopsis thaliana]	gi 332642427,gi 75311602	42,795.50	100.00%	1	0.01%	3.82%
MCR	elongation factor G [Arabidopsis thaliana]	gi 62320893	44,945.90	100.00%	2	0.01%	8.54%
MCR	pyruvate, phosphate dikinase 1 [Arabidopsis thaliana]	gi 145333150,gi 145333164,gi 186511835,gi 193806403,gi	105,140.10	100.00%	13	0.08%	8.88%
MCR	S-methyltetrahydropteroylglutamate--homocysteine methyl	gi 145334507,gi 332005103,gi 332005104,gi 8134566	84,358.40	99.80%	3	0.02%	6.14%
MCR	RecName: Full=Chaperonin CPN60, mitochondrial; AltName: Full=	gi 12644189,gi 15229559,gi 20466256,gi 332643321	61,281.80	97.30%	1	0.01%	2.08%
MCR	RecName: Full=14-3-3-like protein GF14nu; AltName: Full=Ger	gi 3023217	29,825.60	99.70%	6	0.04%	10.20%
MCR	heat shock protein 81-2 [Arabidopsis thaliana]	gi 15241115,gi 2495365,gi 332009328	80,068.00	100.00%	6	0.04%	8.73%
MCR	heat shock protein 70-2 [Arabidopsis thaliana]	gi 15240578,gi 24030296,gi 332008487	76,997.50	100.00%	5	0.03%	8.64%
MCR	proteasome subunit beta type-6 [Arabidopsis thaliana]	gi 15235889,gi 15450505,gi 16323318,gi 21592365,gi 332	25,151.40	99.80%	1	0.01%	5.15%
MCR	cytosolic glyceraldehyde-3-phosphate dehydrogenase [Arabidop	gi 166706,gi 166710	36,986.60	100.00%	8	0.05%	20.10%
MCR	hypothetical protein F16M2.40 - Arabidopsis thaliana	gi 11277801,gi 7523401	32,131.10	98.90%	1	0.01%	3.46%
MCR	phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana]	gi 332642071,gi 73917651	110,164.10	100.00%	11	0.07%	9.19%

Table A2. LC-MS/MS analysis of *Arabidopsis* matched protein and protein cluster report. GSTs further investigated are starred.

MCR	RecName: Full=L-ascorbate peroxidase 2, cytosolic; AltName: f	gi 118572828,gi 257352621,gi 257632823,gi 257643578,gi	27,947.00	86.40%	0	0.00%	0.00%
MCR	RecName: Full=Malate dehydrogenase, cytoplasmic 2	gi 11133446,gi 15239843,gi 332007563	35,674.80	99.60%	2	0.01%	8.43%
MCR	chaperonin-60 alpha subunit [Arabidopsis thaliana]	gi 1223910,gi 15226314,gi 21554572,gi 2506276,gi 25771	62,131.40	100.00%	2	0.01%	4.27%
MCR	ClpC [Arabidopsis thaliana]	gi 2921158	103,458.80	99.80%	4	0.02%	4.85%
MCR	glutathione S-transferase TAU 23 [Arabidopsis thaliana]	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 944	25,682.10	100.00%	4	0.02%	4.09%
MCR	phosphoglycerate kinase, putative [Arabidopsis thaliana]	gi 21536853,gi 298541591	42,148.90	100.00%	3	0.02%	5.99%
MCR	RecName: Full=Heat shock protein 90-1; Short=AtHSP90.1; AltN	gi 26454635,gi 332008861	81,183.70	72.30%	1	0.01%	1.71%
WCS	RecName: Full=Malate dehydrogenase 1, mitochondrial; AltN	gi 11133715,gi 18404382,gi 332194789	35,805.20	98.10%	5	0.03%	6.74%
WCS	phosphoglycerate kinase [Arabidopsis thaliana]	gi 15223484,gi 186491325,gi 332195235,gi 332195236,gi	49,939.50	99.90%	4	0.02%	11.70%
WCS	adenosylhomocysteine 1 [Arabidopsis thaliana]	gi 145333043,gi 15236376,gi 332657945,gi 332657948,gi	53,379.20	90.10%	0	0.00%	0.00%
WCS	fructose-bisphosphate aldolase, class I [Arabidopsis thaliana]	gi 330252066,gi 75313518	42,931.50	100.00%	3	0.02%	3.26%
WCS	D-ribulose-5-phosphate-3-epimerase [Arabidopsis thaliana]	gi 15027997,gi 15240250,gi 20259209,gi 30697525,gi 332	30,008.50	100.00%	5	0.03%	12.50%
WCS	At2g32520 [Arabidopsis thaliana]	gi 110743666,gi 15225693,gi 21554208,gi 30102612,gi 32	25,921.60	99.90%	2	0.01%	8.79%
WCS	putative malate oxidoreductase (malic enzyme) [Arabidopsis t	gi 15225438,gi 20268709,gi 24030498,gi 330251148,gi 45	69,657.30	89.80%	0	0.00%	0.00%
WCS	Unknown protein [Arabidopsis thaliana]	gi 17065446,gi 18399513,gi 20148535,gi 21592994,gi 330	33,797.20	100.00%	1	0.01%	2.67%
WCS	proteasome subunit alpha type-5-B [Arabidopsis thaliana]	gi 12229903,gi 145325425,gi 15220961,gi 15231824,gi 18	25,947.30	100.00%	5	0.03%	14.80%
WCS	proteasome subunit beta type-7-B [Arabidopsis thaliana]	gi 30693626,gi 332007184,gi 332007185,gi 332007186,gi	29,617.60	99.70%	3	0.02%	6.93%
WCS	RecName: Full=Eukaryotic initiation factor 4A-2; Short=eIF-4A-	gi 1170503,gi 1170505,gi 332194952,gi 332194953,gi 332	46,762.90	95.80%	1	0.01%	2.43%
WCS	Elongation factor 1-alpha [Arabidopsis thaliana]	gi 119143,gi 145323788,gi 186532608,gi 332009935,gi 33	49,503.40	100.00%	4	0.02%	4.23%
WCS	RecName: Full=14-3-3-like protein GF14 lambda; AltName: Ful	gi 1345595	27,976.50	81.00%	2	0.01%	4.03%
WCS	transketolase-like protein [Arabidopsis thaliana]	gi 18411711,gi 20268682,gi 21553673,gi 22136900,gi 233	79,841.00	99.80%	2	0.01%	1.35%
WCS	unnamed protein product [Arabidopsis thaliana]	gi 300648936,gi 30698866,gi 332197184,gi 332197186,gi	59,512.40	96.00%	1	0.01%	1.85%
WCS	RecName: Full=Proteasome subunit alpha type-6-B; AltName:	gi 12229902,gi 15224993,gi 300536463,gi 330250885,gi 3	27,350.30	74.80%	0	0.00%	0.00%
WCS	phosphoglycerate kinase 1 [Arabidopsis thaliana]	gi 15230595,gi 332641725,gi 75273138	50,113.10	99.40%	3	0.02%	7.90%
WCS	AT5g47840/MCA23_18 [Arabidopsis thaliana]	gi 10177920,gi 15146328,gi 15238822,gi 22975722,gi 24	31,452.70	100.00%	9	0.05%	7.07%
WCS	phosphoenolpyruvate carboxylase 2 [Arabidopsis thaliana]	gi 240254633,gi 330255051,gi 330255052,gi 73917650	109,757.20	88.20%	1	0.01%	0.94%
WCS	proteasome subunit beta type-3-B [Arabidopsis thaliana]	gi 12744997,gi 17380184,gi 18395025,gi 18411547,gi 213	22,759.80	97.70%	2	0.01%	6.86%
WCS	pyruvate, phosphate dikinase 1 [Arabidopsis thaliana]	gi 145333150,gi 145333164,gi 186511835,gi 193806403,gi	105,140.10	100.00%	12	0.07%	4.08%
WCS	S-methyltetrahydropteroyltriglutamate--homocysteine methy	gi 145334507,gi 332005103,gi 332005104,gi 8134566	84,358.40	87.50%	1	0.01%	1.31%
WCS	RecName: Full=14-3-3-like protein GF14 nu; AltName: Full=Ger	gi 3023217	29,825.60	87.00%	2	0.01%	3.77%
WCS	proteasome subunit beta type-6 [Arabidopsis thaliana]	gi 15235889,gi 15450505,gi 16323318,gi 21592365,gi 332	25,151.40	64.30%	0	0.00%	0.00%
WCS	cytosolic glyceraldehyde-3-phosphate dehydrogenase [Arabido	gi 166706,gi 166710	36,986.60	89.80%	0	0.00%	0.00%
WCS	hypothetical protein F16M2.40 - Arabidopsis thaliana	gi 11277801,gi 7523401	32,131.10	98.80%	1	0.01%	3.46%
WCS	RecName: Full=Glyceraldehyde-3-phosphate dehydrogenase f	gi 20455491,gi 332193812	47,659.30	95.40%	1	0.01%	2.24%
WCS	phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana]	gi 332642071,gi 73917651	110,164.10	99.80%	3	0.02%	2.38%
WCS	RecName: Full=L-ascorbate peroxidase 2, cytosolic; AltName: f	gi 118572828,gi 257352621,gi 257632823,gi 257643578,gi	27,947.00	62.70%	0	0.00%	0.00%
WCS	RecName: Full=Malate dehydrogenase, cytoplasmic 2	gi 11133446,gi 15239843,gi 332007563	35,674.80	66.20%	1	0.01%	4.82%
WCS	ClpC [Arabidopsis thaliana]	gi 2921158	103,458.80	57.60%	0	0.00%	0.00%
WCS	glutathione S-transferase TAU 23 [Arabidopsis thaliana]	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 944	25,682.10	99.90%	11	0.06%	4.09%
WCS	phosphoglycerate kinase, putative [Arabidopsis thaliana]	gi 21536853,gi 298541591	42,148.90	99.80%	2	0.01%	3.24%

Table A2. Continued

Biological sample name	Protein Cluster	Protein accession numbers	Protein identification probability	Number of total spectra	Percentage of total spectra	Maximum sequence coverage
ACR	Cluster of DING protein [Arabidopsis thaliana] (gi 53854257)	gi 53854257	100.00%	2	0.01%	49.10%
ACR	★ Cluster of glutathione S-transferase TAU 23 [Arabidopsis thaliana] (gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	3	0.02%	4.09%
ACR	Cluster of unnamed protein product [Arabidopsis thaliana] (gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	5	0.03%	12.00%
ACR	Cluster of RecName: Full=Elongation factor 1-alpha; Short=EF-1-alpha; /gi 119143,gi 145323788,gi 186532608,gi 332009935	gi 119143,gi 145323788,gi 186532608,gi 332009935	100.00%	2	0.01%	4.68%
ACR	Cluster of phosphoglycerate kinase [Arabidopsis thaliana] (gi 15223484,gi 15230595,gi 186491325,gi 33219523)	gi 15223484,gi 15230595,gi 186491325,gi 33219523	100.00%	4	0.03%	12.10%
ACR	Cluster of pyruvate, phosphate dikinase 1 [Arabidopsis thaliana] (gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806)	gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806	100.00%	6	0.04%	4.50%
ACR	Cluster of phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana] (gi 240254633,gi 24636264,gi 330255051,gi 3302550)	gi 240254633,gi 24636264,gi 330255051,gi 3302550	100.00%	2	0.01%	2.38%
ACR	Cluster of fructose-bisphosphate aldolase, class I [Arabidopsis thaliana] (gi 330250312,gi 330252066,gi 75268018,gi 7531351)	gi 330250312,gi 330252066,gi 75268018,gi 7531351	100.00%	3	0.02%	7.02%
ACR	Cluster of RecName: Full=Malate dehydrogenase 1, mitochondrial; AltName: F_0gi 11133715,gi 145332399,gi 15232468,gi 18404382	gi 11133715,gi 145332399,gi 15232468,gi 18404382	100.00%	3	0.02%	11.10%
ACR	Cluster of RecName: Full=14-3-3-like protein GF14 lambda; AltName: F_0gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 53854257	gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 53854257	100.00%	1	0.01%	4.84%
MCR	Cluster of DING protein [Arabidopsis thaliana] (gi 53854257)	gi 53854257	100.00%	2	0.01%	54.70%
MCR	★ Cluster of glutathione S-transferase TAU 23 [Arabidopsis thaliana] (gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	4	0.02%	4.09%
MCR	Cluster of unnamed protein product [Arabidopsis thaliana] (gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	4	0.02%	12.00%
MCR	Cluster of RecName: Full=Proteasome subunit alpha type-5-A; AltName: F_0gi 12229903,gi 145325425,gi 15220961,gi 15231824	gi 12229903,gi 145325425,gi 15220961,gi 15231824	100.00%	2	0.01%	10.50%
MCR	Cluster of phosphoglycerate kinase, putative [Arabidopsis thaliana] (gi 21536853,gi 298541591)	gi 21536853,gi 298541591	100.00%	3	0.02%	5.99%
MCR	Cluster of cystolic glyceraldehyde-3-phosphate dehydrogenase [Arabidopsis thaliana] (gi 166706,gi 166710)	gi 166706,gi 166710	100.00%	8	0.05%	20.10%
MCR	Cluster of thiol protease aleurain-like protein [Arabidopsis thaliana] (gi 332644500,gi 332644501,gi 73622182,gi 7931427)	gi 332644500,gi 332644501,gi 73622182,gi 7931427	100.00%	2	0.01%	7.26%
MCR	Cluster of chaperonin-60 alpha subunit [Arabidopsis thaliana] (gi 1223910,gi 15226314,gi 21554572,gi 2506276,gi 1170503,gi 1170505,gi 332194952,gi 332194953,gi 15240578)	gi 1223910,gi 15226314,gi 21554572,gi 2506276,gi 1170503,gi 1170505,gi 332194952,gi 332194953,gi 15240578	100.00%	2	0.01%	5.83%
MCR	Cluster of heat shock protein 70-2 [Arabidopsis thaliana] (gi 15240578)	gi 15240578	100.00%	5	0.03%	8.64%
MCR	Cluster of 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase [Arabidopsis thaliana] (gi 145334507,gi 145334531,gi 186509731,gi 3320005)	gi 145334507,gi 145334531,gi 186509731,gi 3320005	100.00%	3	0.02%	6.14%
MCR	Cluster of heat shock protein 81-2 [Arabidopsis thaliana] (gi 15241115,gi 15241115,gi 2495365,gi 26454635,gi 332008861,gi 332641697,gi 12643245,gi 12643273,gi 12644165,gi 18206367,gi 186478427,gi 20455491,gi 332190823,gi 3321908)	gi 15241115,gi 15241115,gi 2495365,gi 26454635,gi 332008861,gi 332641697,gi 12643245,gi 12643273,gi 12644165,gi 18206367,gi 186478427,gi 20455491,gi 332190823,gi 3321908	100.00%	6	0.04%	8.73%
MCR	Cluster of heat shock protein 70-4 [Arabidopsis thaliana] (gi 332641697,gi 12643245,gi 12643273,gi 12644165,gi 18206367,gi 186478427,gi 20455491,gi 332190823,gi 3321908)	gi 332641697,gi 12643245,gi 12643273,gi 12644165,gi 18206367,gi 186478427,gi 20455491,gi 332190823,gi 3321908	100.00%	18	0.11%	17.70%
MCR	Cluster of glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphoglycerate kinase) [Arabidopsis thaliana] (gi 186478427,gi 20455491,gi 332190823,gi 3321908)	gi 186478427,gi 20455491,gi 332190823,gi 3321908	100.00%	7	0.04%	14.30%
MCR	Cluster of RecName: Full=Elongation factor 1-alpha; Short=EF-1-alpha; /gi 119143,gi 145323788,gi 186532608,gi 332009935	gi 119143,gi 145323788,gi 186532608,gi 332009935	100.00%	3	0.02%	6.46%
MCR	Cluster of phosphoglycerate kinase [Arabidopsis thaliana] (gi 15223484,gi 15230595,gi 186491325,gi 33219523)	gi 15223484,gi 15230595,gi 186491325,gi 33219523	100.00%	3	0.02%	8.58%
MCR	Cluster of Transketolase [Arabidopsis thaliana] (gi 18411711,gi 20268682,gi 21553673,gi 22136900,gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806)	gi 18411711,gi 20268682,gi 21553673,gi 22136900,gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806	100.00%	4	0.02%	9.45%
MCR	Cluster of pyruvate, phosphate dikinase 1 [Arabidopsis thaliana] (gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806)	gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806	100.00%	13	0.08%	8.68%
MCR	Cluster of phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana] (gi 240254633,gi 24636264,gi 330255051,gi 3302550)	gi 240254633,gi 24636264,gi 330255051,gi 3302550	100.00%	13	0.08%	9.19%
MCR	Cluster of RecName: Full=ATP synthase subunit beta, chloroplastic; AltName: F_0gi 6686269,gi 7525040	gi 6686269,gi 7525040	100.00%	15	0.09%	27.50%
MCR	Cluster of adenosylhomocysteine lyase 1 [Arabidopsis thaliana] (gi 145333043,gi 15236376,gi 332657945,gi 3326579)	gi 145333043,gi 15236376,gi 332657945,gi 3326579	100.00%	4	0.02%	8.45%
MCR	Cluster of peptide methionine sulfoxide reductase A2 [Arabidopsis thaliana] (gi 15240795,gi 332003778,gi 38454142,gi 46402480)	gi 15240795,gi 332003778,gi 38454142,gi 46402480	99.80%	2	0.01%	10.10%
MCR	Cluster of fructose-bisphosphate aldolase, class I [Arabidopsis thaliana] (gi 330250312,gi 330252066,gi 75268018,gi 7531351)	gi 330250312,gi 330252066,gi 75268018,gi 7531351	100.00%	3	0.02%	10.80%
MCR	Cluster of peptide methionine sulfoxide reductase A3 [Arabidopsis thaliana] (gi 15240796,gi 332003779,gi 75264361)	gi 15240796,gi 332003779,gi 75264361	100.00%	2	0.01%	10.90%
MCR	Cluster of ClpC [Arabidopsis thaliana] (gi 2921158,gi 1169544,gi 18423233,gi 186510816,gi 2921158,gi 18391066,gi 19548023,gi 2160168,gi 30316342,gi 11133715,gi 145332399,gi 15232468,gi 18404382)	gi 2921158,gi 1169544,gi 18423233,gi 186510816,gi 2921158,gi 18391066,gi 19548023,gi 2160168,gi 30316342,gi 11133715,gi 145332399,gi 15232468,gi 18404382	100.00%	6	0.04%	5.67%
MCR	Cluster of 2,3-bisphosphoglycerate-independent phosphoglycerate kinase [Arabidopsis thaliana] (gi 18391066,gi 19548023,gi 2160168,gi 30316342,gi 11133715,gi 145332399,gi 15232468,gi 18404382)	gi 18391066,gi 19548023,gi 2160168,gi 30316342,gi 11133715,gi 145332399,gi 15232468,gi 18404382	100.00%	2	0.01%	3.05%
MCR	Cluster of RecName: Full=Malate dehydrogenase 1, mitochondrial; AltName: F_0gi 11133715,gi 145332399,gi 15232468,gi 18404382	gi 11133715,gi 145332399,gi 15232468,gi 18404382	100.00%	5	0.03%	11.40%
MCR	Cluster of RecName: Full=Malate dehydrogenase, cytoplasmic 2 (gi 11133446,gi 11133509,gi 15219721,gi 15239843,gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 62320893)	gi 11133446,gi 11133509,gi 15219721,gi 15239843,gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 62320893	100.00%	2	0.01%	8.43%
MCR	Cluster of RecName: Full=14-3-3-like protein GF14 lambda; AltName: F_0gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 62320893)	gi 12230867,gi 1345595,gi 152112421,gi 1702987,gi 62320893	100.00%	6	0.04%	10.40%
MCR	Cluster of elongation factor G [Arabidopsis thaliana] (gi 62320893)	gi 62320893	100.00%	2	0.01%	8.54%
MCR2	Cluster of RecName: Full=Probable phosphoglucomutase, cytoplasmic (gi 12585324,gi 238479031,gi 238479033,gi 3321969)	gi 12585324,gi 238479031,gi 238479033,gi 3321969	100.00%	2	0.01%	3.59%
WCS	Cluster of putative ribulose-5-phosphate-3-epimerase [Arabidopsis thaliana] (gi 15027997,gi 15240250,gi 20259209,gi 30697525,gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 15027997,gi 15240250,gi 20259209,gi 30697525,gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	5	0.03%	12.50%
WCS	★ Cluster of glutathione S-transferase TAU 23 [Arabidopsis thaliana] (gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 15218301,gi 21553948,gi 332197974,gi 8052536,gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	11	0.06%	4.09%
WCS	Cluster of unnamed protein product [Arabidopsis thaliana] (gi 10177920,gi 15146328,gi 15238822,gi 227295722)	gi 10177920,gi 15146328,gi 15238822,gi 227295722	100.00%	9	0.05%	7.07%
WCS	Cluster of putative carboxymethylenebutenolidase [Arabidopsis thaliana] (gi 110743666,gi 15225693,gi 21554208,gi 30102612)	gi 110743666,gi 15225693,gi 21554208,gi 30102612	100.00%	2	0.01%	8.79%
WCS	Cluster of RecName: Full=Proteasome subunit alpha type-5-A; AltName: F_0gi 12229903,gi 145325425,gi 15220961,gi 15231824	gi 12229903,gi 145325425,gi 15220961,gi 15231824	100.00%	5	0.03%	14.80%
WCS	Cluster of auxin-induced atb2-like protein [Arabidopsis thaliana] (gi 30696459,gi 332195599,gi 332195603,gi 3321956)	gi 30696459,gi 332195599,gi 332195603,gi 3321956	100.00%	2	0.01%	8.18%
WCS	★ Cluster of RecName: Full=Glutathione S-transferase 16; AltName: Full=gi 12230166,gi 148887379,gi 15227063,gi 18655397	gi 12230166,gi 148887379,gi 15227063,gi 18655397	100.00%	15	0.09%	5.19%
WCS	Cluster of proteasome subunit beta type-7-B [Arabidopsis thaliana] (gi 30693626,gi 332007184,gi 332007185,gi 3320071)	gi 30693626,gi 332007184,gi 332007185,gi 3320071	99.90%	3	0.02%	6.93%
WCS	Cluster of RecName: Full=Elongation factor 1-alpha; Short=EF-1-alpha; /gi 119143,gi 145323788,gi 186532608,gi 332009935	gi 119143,gi 145323788,gi 186532608,gi 332009935	100.00%	4	0.02%	4.23%
WCS	Cluster of phosphoglycerate kinase [Arabidopsis thaliana] (gi 15223484,gi 15230595,gi 186491325,gi 33219523)	gi 15223484,gi 15230595,gi 186491325,gi 33219523	100.00%	4	0.02%	11.70%
WCS	Cluster of pyruvate, phosphate dikinase 1 [Arabidopsis thaliana] (gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806)	gi 14,gi 145333150,gi 145333164,gi 186511835,gi 193806	100.00%	12	0.07%	4.08%
WCS	Cluster of phosphoenolpyruvate carboxylase 3 [Arabidopsis thaliana] (gi 240254633,gi 24636264,gi 330255051,gi 3302550)	gi 240254633,gi 24636264,gi 330255051,gi 3302550	100.00%	4	0.02%	3.31%
WCS	Cluster of RecName: Full=Malate dehydrogenase 1, mitochondrial; AltName: F_0gi 11133715,gi 145332399,gi 15232468,gi 18404382	gi 11133715,gi 145332399,gi 15232468,gi 18404382	100.00%	5	0.03%	7.28%

Table A2. Continued

Gene designation	Gene bank match (Arabidopsis)	Size (bp)	Matching waterhemp contig sequence (ACR)
β Tubulin	gi 13926187	346	TACCATGGACAGTATTCGGGCCGGCCCTTATGGCCAGATTTTCGCCCGGATAACTTTGTTTTGGACAATC TGGGGCAGGAAATAATTGGGCTAAAGGTCATTATACTGAAGGTGCTGAGTTGATTGATTCGGTCTAGATG TTGTTGAAAAAGAGCTGAGAATTGTGATTGTATGCAAGGGTTTCAAGTTTGTCATTCTGGGAGGAGGA ACAGGATCTGGAATGGGAACATTGCTGATTTCAAAGATACGAGAGGAGTATCCAGATCGAATGATGATGA CATTTTCGGTTTTCCCTTCGCCGAAGGTTTCAGATACTGTGGTTGAGCCTTACAATGCTACA
<i>ArGSTU1</i>	GI:15220041	590	TATGAGAATAGTAGGAGAATTTGGACTACAAAAGGTGAAGAACAAGAGGCCAGCAAGAAGGAATTCATA GGCCATTTCAAATTATTAGAGCAACAATTAGGAAACAAGCCATATTTGGAGGGGATTCTTTGGGTTTGT GATGTTAGTCTCATCCATTCTATTGTTGTTTTATCTTGGGAAAACTAGGCAACATTAGTATTGGTGAG TCATGCCCTAAGCTTATTGAATGGGCTAAAAGATGCATGGAAAGAGATAGTGTGCTAAGTCTCTCAAGA TGAGAAGAAGATCTATGAGTTCTATTGGAGTTGAAGAAGGCACTTGGTATCGAGTAGGGGAAACGCGGT ACCTCTGTTCTAAAGATATGTTCCATCTTTTATTCATATGAGTTTATCATCGTCATCTAATGAGTTTCGCTA GAGGCAGGGTCTGGTTATTCATATGAGTCTATTAGTCTAATCAACACGAATAGATCGTTGGAACGGTTTA TGTTGTATGGTAGATTGAAGTATGGGTTGTGTTGTATGGTAGATTGAAGTATGGGTTGAGTACTTGAGTTT TTTTTTGTTGTTGATNTG
<i>ArGSTU2</i>	GI:15220041	439	TAATGTTCTATCCCAAGATAAACTCAGAAAGTTTATCCATTGAGGAATAGATGTGTTAACAGAATGTCTA AGTACACATGTTTTACCCCAAGCAATGATTTTGGGACACTCTAATTCATATTTAAACTTCCACAAGTTTCAT AAACTTGAAACCAAGCTGTAAATGGAATTAGGGCAATGTCCATATACCCAAATTTATCTCCTTCAAAAAG ACTTATCTCCCAACACTTGTCTAGTGTCTTACACATTCTATGAACCTGTTTTGCTTCTCTCTTTTGC TATCCTTTGTTCTCATAACTTATTGCTTGCTTCAAAAAAAGTATGTAATCCAACTGAGCCCAAGATCTT GATTGTGCCCTCTGTTAAGGATCAGAAGACAGAAAATGAGTATCCCTTTCCAAATTCATCAATATACCTC AC
<i>ArGSTU3</i>	GI:15220041	736	GCCACAAAAGAATAATCTTTATTAATTATTAATTATAACACAACATCCAATCATAATTGAAAAAATAATACA ACAATAACAACTTTGTGTACTGGCAGTACTTATTTTCAGTATGACAACTTCAATAAAGTTCCAGAACAC AAAAATCATAGAAAAATAAGGCACTTCAACAGGTATATCAAGGTAGACCTTTATCTTGCAATTCACGAA GAGTCTACTTCTAAATCAACCCGTAAACCTTTCAACACACAACAACACTTAAACCGAAGCATCATAAAAA CACAAACATAGAAACCATACTTAATTCCTCCGTTCCATTATTTACATCTGATATGTAACATTGTCACTATC AGGTACAAGTATTTCAAGACGGAGGTAAATATGATCTCTATTGTTCCAAAAAATCTTTCTCATGGCCATA CACTTTTGGCAAGAAGGCTAGCCAAAATTTCTCACGCCAAGCATTGACATGAGGGCGTTCTCGAACATC TTTATAATTTTGTACCACTCAATGTTTGAAGGCGGGCATGTGATGAACATCTCTAGACTAAACGTATCA CCCGCTACGTACTTATTCTCGACAAATACGCTTCGTATATGTCGAGGACCTTTCGGAATTTAGGCTCGTATT CATCTACTATTGCATCGTCAAGCCAAGCCCATGTAACCTTGTGAAGCAAGCTCATTGCGAGTTTAAGGG CGATCGACTCGAAT
<i>ArGSTF1</i>	GI:330250548	305	CGCGGGGCCACCCTCCGGTTTTGGCAGCTGCTAATGAGAAAGAGCTTCAATCCAATTGGTTGATGTGCA TTTGTCTTCCGGGGCTAACAACAACCATCTTTCTTTCTTTAATCCTTTTGGTCAAAATCCCTGCTTTTGAAG ATGGAGATCTCACTTTATTGAATCAAGGGCAATAACAAAATACATAGCACACCTTATGAAGAACAGGA ACACCTCTTACATACAAAACAGGGAAAGAAATGGCAGATCAAGCCATTTGGATGGAAGTAGAGGCTCACC AATTTGATCCAGTAGTAAGT
<i>ArGSTF2</i>	GI:330250548	228	CCGGATGCACCCAACGTGTATTAGCCGCTGCTCATGAGAAGAATCTTGATTATGAATTGTTTATTGTTGACT TGAGGAAACACCAACAAAAGGAACCATCTTTCTCTCCCTCAATCCGTTTGGTCAAGTCCCTGTTTCCAAG ATGGAGATCTCAAAATTGATTGAATCAAGGGCAATAACAAGATACATAGCATACACTTACGAAGGCCAAGG AACACCCCTTACTT

Table A3. Gene identities and contig sequences. Genes listed represent LC-MS identities matched with contigs from the sequenced ACR transcriptome (Lee et al., 2009).

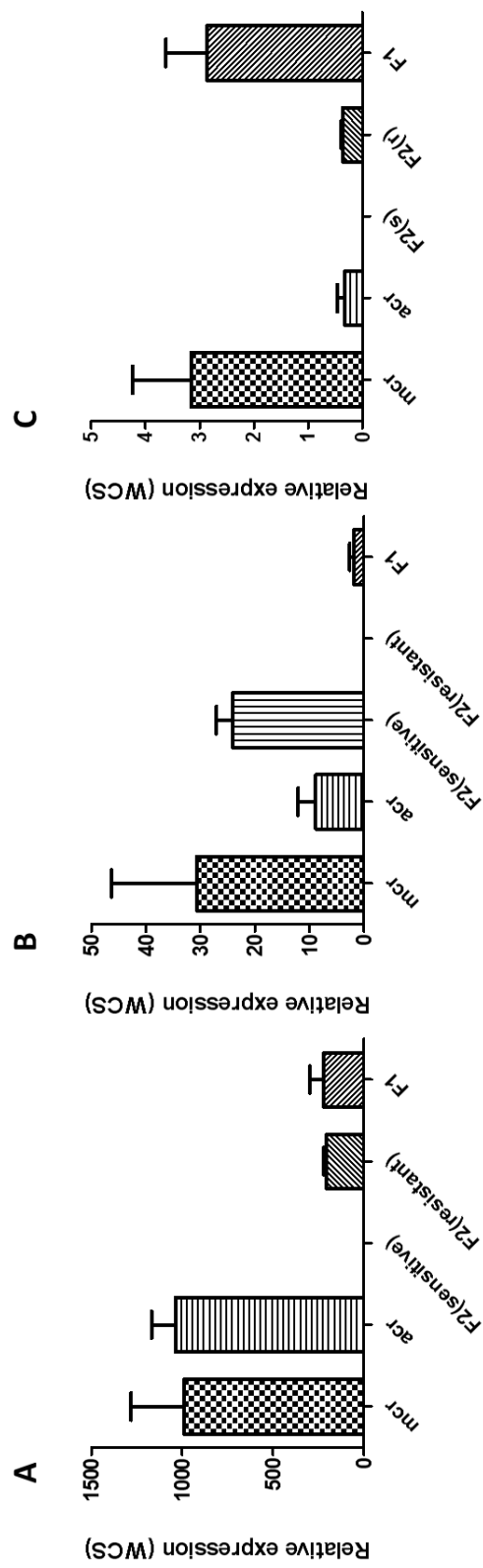


Figure A1. Expression of candidate GSTs in MCR, ACR, F₁ hybrid (MCR x WCS) and F₂ lines (MCR x WCS) relative to WCS. qRT-PCR assays were conducted for each target GST using primers from Table 3. Dissociation curves for every reaction were analyzed to ensure only one replicon was amplified. Expression of *ArGSTF2* (A), *ArGSTU1* (B), and *ArGSTU2* (C) in MCR and ACR was calculated relative to expression in WCS using the 2- $\Delta\Delta$ CT method (Livak 2001), with expression of beta-tubulin (Table A3) as a constitutive control gene. Data represent the mean of three biological replications with three technical replications each. Bars represent the standard error of the mean. McLean County (atrazine-resistant [MCR]), Adams County (atrazine-resistant [ACR]), and Wayne County (atrazine-sensitive [WCS]) Illinois, as described in (Hausman et al. 2011; Ma et al. 2013).