DISSERTATION

INVOLVEMENT OF *CYP72A219* IN HERBICIDE-RESISTANT PALMER AMARANTH AND THE ROLE OF P450 REDUCTASE IN THE MECHANISM OF METABOLIC RESISTANCE

Submitted by

Carlos A. G. Rigon

Department of Agricultural Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2023

Doctoral Committee:

Advisor: Todd A. Gaines
Co-Advisor: Franck E. Dayan

Roland Beffa
Christie Peebles

ABSTRACT

INVOLVEMENT OF *CYP72A219* IN HERBICIDE-RESISTANT PALMER AMARANTH AND THE ROLE OF P450 REDUCTASE IN THE MECHANISM OF METABOLIC RESISTANCE

Herbicide resistance in weeds poses a major challenge to modern agriculture worldwide, impacting effective weed control strategies. Metabolic resistance stands out as the major and more complex resistance mechanism due to its ability to metabolize a wide range of herbicides within weed species. Metabolic resistance involves herbicide metabolism through three key phases: activation, conjugation, and sequestration. These phases involve the action of important enzymes such as cytochrome P450 monooxygenases, glutathione *S*-transferases, and ABC transporters. Metabolic resistance mechanisms have gained prominence in the past decade, posing significant challenges to sustainable agriculture and weed management practices.

Amaranthus palmeri (Palmer amaranth) one of the most troublesome weeds globally has evolved metabolic resistance to HPPD inhibitor tembotrione. Understanding and addressing the mechanism are crucial for developing effective strategies to combat herbicide resistance and ensure global crop production. In the present study, four upregulated P450 genes were identified in HPPD-resistant Palmer amaranth from Nebraska (NER), a troublesome weed species. Among these genes, *CYP72A219_4284* demonstrated the ability to deactivate the herbicide tembotrione in a heterologous system. This gene was also upregulated in metabolic HPPD-resistant Palmer amaranth plants from different fields across the United States, indicating its involvement in conferring herbicide resistance. Our study also investigated the regulation of these resistance

genes, including the promoter sequences and transcription factors involved. Additionally, quantitative trait loci associated with herbicide resistance were identified. This work represents the first identification and validation of genes responsible for herbicide metabolism in Palmer amaranth. Validation of the metabolic resistant gene and the exploration of regulatory mechanisms contribute to a better understanding of metabolic herbicide resistance in weeds, facilitating the development of effective weed management strategies.

Cytochrome P450 reductase (CPR), an essential enzyme localized in the endoplasmic reticulum, provides electrons for P450 enzymes during monooxygenase reactions. The transfer of electrons from NADPH to the P450 active site occurs through a complex CPR:P450 interaction. Despite the numerous P450 genes in plant genomes, CPR genes are limited, typically consisting of two or three copies. In Arabidopsis, the two CPR genes, *ATR1* and *ATR2*, have distinct roles in primary and inducible metabolism, respectively. Our study investigated the function of *ATR1* and *ATR2* in transgenic Arabidopsis plants overexpressing the *CYP81A12*, which is known to metabolize a wide range of herbicides. The hypothesis was that silencing these *ATR1* or *ATR2* genes would lead to a reduction of P450 activity involved in herbicide metabolism. *ATR1* predominantly transfers electrons to *CYP81A12*, as knocking down *ATR1* led to a significant reduction in herbicide resistance. Knockouts of the *ATR2* gene also resulted in decreased herbicide resistance, although the effect was less pronounced. Variation in the number and function of CPR genes among different weed species suggests diverse genetic pressures and potential targets for herbicide resistance management. Inhibition of CPR activity could be a promising approach to restore herbicide effectiveness against metabolic herbicide-resistant weeds. This is the first study to our knowledge that explores the involvement of CPR genes in herbicide resistance in weeds, providing valuable insights into their crucial role. The findings

significantly advance our understanding of the mechanisms underlying CPR-mediated herbicide resistance and offer potential targets for the development of effective weed management strategies.

ACKNOWLEDGEMENTS

I am humbled and deeply grateful as I take this moment to extend my heartfelt appreciation to the many individuals who have played a significant role in shaping my journey. First and foremost, I offer my sincerest thanks to God, for the gift of life and the abundance of opportunities that have come my way. I am truly blessed to have been surrounded by such wonderful people and experiences.

To my beloved wife, Joice A Rigon, your unwavering support and belief in me have been a constant source of strength and motivation. Your presence in my life has made all the difference, and I am forever grateful for your love. Thank you so much!

To my family back in Brazil, my dear mother Rosane Gonsiorkiewicz, my father João Rigon, and my brother João Paulo G Rigon, thank you for your endless love and encouragement. Your support has been instrumental in shaping my character and helping me become the person I am today. I am aware that we have missed out on many precious moments together due to my career aspirations, and I want you to know that your sacrifice and encouragement mean the world to me.

I must also express my gratitude to the educational institutions and teachers who have contributed to my growth and knowledge. The education I received has been the foundation upon which I have built my journey, both professionally and personally.

To my advisors, Todd Gaines and Franck Dayan, I am deeply thankful for your confidence in me and for pushing me towards my goals. Your guidance and support have been invaluable, and I have learned so much from both of you. Choosing to pursue my PhD under your mentorship has been one of the best decisions of my life, as the exceptional teaching and

mentorship I received have undoubtedly prepared me for my professional career ahead. I extend my gratitude to the members of my committee, Rolland Beffa and Christie Peebles, for their insightful discussions and guidance throughout the research process. "*If I have seen further, it is by standing on the shoulders of giants*" (Issac Newton, 1675).

I am incredibly grateful to my lab mates who have been with me on this research journey for varying periods, ranging from a few months to over two years. Each one of you has played a vital role in contributing to the success and progress of this project, and I extend my heartfelt appreciation to all.

To Alan and Cathy Hendrickson, you have been more than friends; you have been family for Joice and me. Your love and support have touched our lives deeply, and we are forever grateful for the blessings of meeting you. The friendship family will stay with us forever.

To the wonderful people from the churches, including Randy and Jan Babcock, Steve, Mel, Bonnie, John, and others, I thank you for your compassion and care for international students. Your kindness has been a tremendous source of comfort and strength.

I extend my gratitude to the Weed Resistance Competence Center at Bayer AG and Cotton Inc. for their generous funding support for this research. A special thanks goes to Anita Küpper from Bayer AG for her unwavering support, friendship, and engaging scientific discussions. The opportunity to intern in Germany with her group was an invaluable experience, and I am truly thankful for the chance to collaborate and contribute to the research. I would also like to express my appreciation to Falco Peter, Veronika, Thomas, Simon, Julia, Jagoda, Evlampia, Alberto Collavo, Bodo Peter, and all others who supported and contributed to the research during my time in Germany. Working with each one of you was a memorable and enriching experience. Thank you all for making this journey unforgettable.

To all those whose names may not be mentioned but have contributed to my journey in their own ways, thank you for being part of my life and for enriching it with your presence and support.

As I move forward in life, I carry with me the love and encouragement of all these wonderful individuals. May their kindness and generosity continue to inspire me to spread love, compassion, and understanding wherever I go.

With gratitude,

Carlos AG Rigon

"Throughout life, you encounter numerous judgments from other, but only a few who are genuinely willing to invest the hard work needed to create a meaningful impact."

TABLE OF CONTENTS

1. INTRODUCTION

The world is currently witnessing a significant rise in food demand, primarily driven by the escalating global population[1]. It is estimated that by 2050, the world population will reach approximately 9.1 billion individuals, compared to the current 8 billion[1]. Pests pose a major threat to the goal of increasing crop production by causing significant reductions in yield[2]. Among the various crop protection factors, weeds stand out as the most prominent pests due to their competition with crops for essential resources such as light, nutrients, and water[3]. As a consequence, weed infestation can lead to a considerable reduction of around 34% in crop yield[2, 4]. The magnitude of this issue amplifies when weeds are resistant to herbicides, creating a major hurdle for modern agriculture worldwide and compromising the effectiveness of weed control strategies[5].

Weeds play a crucial role in agricultural systems, and their ability to evolve resistance poses a significant challenge. The extensive use of herbicides has led to the selection of herbicide-resistant biotypes from various weed species[6]. Currently, there are over 522 documented cases of resistant weeds worldwide[7]. Among the most prevalent resistance cases are those targeting acetolactate synthase (ALS), Photosystem II (PSII) 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), acetyl Coa-carboxylase (ACCase), auxinic herbicides, PSI, and microtubule herbicides[7]. Notably, *Lolium rigidum*, *Echinochloa crus-galli*, *Poa annua*, *Amaranthus palmeri*, and *Eleusine indica* are among the key weed species that have evolved resistance to multiple herbicides, however, 269 species have evolved resistance including 154 dicots and 115 monocots.[6, 7]. The rapid evolution of herbicide resistance across multiple herbicides and weed species highlights the urgent need for effective strategies to manage these resistant weeds and ensure sustainable agricultural practices.

Herbicide resistance in weeds arises through diverse mechanisms that enable their survival and reproduction despite herbicide treatments. These mechanisms of resistance are categorized into target-site (TSR) and non-target site resistance (NTSR)[6, 8, 9]. TSR involves genetic mutations occurring within the herbicide's target site DNA. This common mechanism caused by mutations can alter the structure of the target protein, decreasing the binding affinity of the herbicide[6]. In some cases, the target gene may undergo amplification or overexpression, leading to an increased production of the target protein and reducing the ratio of herbicide molecule and target protein. Some examples are the *EPSPS* in glyphosate resistant *Amaranthus palmeri*[10] and *ACCAse* in resistant *Digitaria sanguinalis*[11]. On the other hand, NTSR encompasses a range of mechanisms that operate outside the direct target site of the herbicide[8, 12]. Reduced translocation or absorption of the herbicide occurs when the weed develops traits that limit the movement of the herbicide within its tissues[6, 13]. Sequestration is when a weed actively isolates the herbicide within specialized cellular compartments, minimizing its interaction with critical sites of action[14]. Additionally, weeds can evolve enhanced metabolic detoxification mechanisms, such as increased enzymatic activity, to break down or inactivate the herbicide before it can exert its intended effect[8].

Among the NTSR, enhanced metabolic resistance poses a significant threat to weed management efforts[8]. This form of resistance is particularly challenging because of its unpredictable nature and the ability of weeds to metabolize herbicides from different chemical families[15]. This mechanism involves three main phases: Phase I involves oxidative reactions. Cytochrome P450 monooxygenases (P450s) are key enzymes in this phase and catalyze the introduction of polar functional groups (e.g., hydroxyl, carboxyl) into the herbicide molecule, making it more reactive and susceptible to further modification[8, 16]. Phase II known as

conjugation, involves the attachment of a polar molecule (e.g., sugars, amino acids, glutathione) to the herbicide. This conjugation process is mediated by various enzymes, including UDP-glycosyltransferases (UGTs), glutathione *S*-transferases (GSTs), and amino acid transferases[8, 17]. Phase III encompasses processes related to the transport and compartmentalization of herbicide metabolites[8]. ATP-binding cassette (ABC) transporters are integral membrane proteins that play a vital role in Phase III of the metabolism by sequestrating the herbicide in the vacuole or other cellular compartments, limiting their interaction with target sites and minimizing their phytotoxic effects[18-20].

Numerous P450 enzymes have been identified in crops, and recently, a significant number of genes have been discovered in weeds that are responsible for herbicide metabolism. Noteworthy examples include various P450s found in *Echinochloa phyllopogon*, such as *CYP81A12*, *CYP81A21, CYP81A15*, and *CYP81A14*[21, 22]. These enzymes are capable of metabolizing herbicides from more than five different modes of action. Additionally, *CYP81A69* and *CYP81A70* from *Cynodon dactylon* have been identified as enzymes involved in metabolizing ALS inhibitors and auxinic herbicides[23]. Furthermore, *CYP81A10v7* from *Lolium rigidum* plays a role in metabolizing ACCase-, ALS-, PSII-, HPPD-, and tubulin-inhibitors[24]. Dicot weeds have a limited number of identified genes associated with herbicide resistance. Notable examples include *DsCYP77B34* from *Descurainia sophia*, which provides resistance to ALS-, PPO-, VLCFA-, and PSII-inhibitors[25]. Additionally, *RrCYP704C1* or *RrCYP709B1* from *Raphanus raphanistrum* confer resistance specifically against HPPD-inhibitors[26]. These identified genes highlight the capacity of dicot weeds to develop resistance mechanisms against certain herbicide classes by enhanced metabolism, which was not previously considered.

Every P450 in plants necessitates the involvement of a partner P450 reductase for its proper functioning, called cytochrome P450 reductase (CPR)[27]. Each reaction of P450 monooxygenase requires subsequent reactivation through CPR. While the plant kingdom boasts numerous P450 monooxygenases, they universally rely on one, two or three P450 reductases[27-29]. In Arabidopsis, two CPR genes exist, known as *Arabidopsis thaliana* cytochrome P450 reductase 1 (*ATR1*) and 2 (*ATR2*) [30, 31]. These isoforms are categorized into two distinct classes, each serving different functions. Class I comprises *ATR1*, which plays a crucial role in primary and basal metabolic processes. On the other hand, Class II encompasses *ATR2*, which demonstrates inducibility in response to environmental stimuli such as wounding, pathogen attacks, or exposure to high light intensity[32, 33]. The reliance of P450 enzymes on CPRs presents an excellent opportunity for the development of innovative synergistic chemistry. These potential advancements hold the promise of enhancing the effectiveness of herbicides, opening new avenues for improved weed control and crop protection.

HPPD inhibitors have emerged as effective herbicides in weed management strategies for crops such as corn, barley, oat, rice, and wheat[34]. To date, only three weed species, namely *Amaranthus palmeri* (Palmer amaranth), *Amaranthus tuberculatus* (common waterhemp), and *Raphanus raphanistrum* (wild radish), have been identified as resistant to HPPD inhibitors[35-37]. HPPD stands for 4-hydroxyphenylpyruvate dioxygenase, an essential enzyme involved in the biosynthesis of carotenoids, a class of pigments crucial for plant growth and development. By inhibiting the activity of HPPD, these herbicides disrupt the production of carotenoids, leading to the impairment of chlorophyll synthesis, then "bleaching", and ultimately causing weed death[38, 39]. They are characterized by low application rates and can be applied either pre- or post-emergence, providing flexibility in weed control timing.[38, 40, 41]. Moreover, HPPD inhibitors offer

effective control against challenging and hard-to-manage weed species, including those that have developed resistance to other herbicide classes, such as glyphosate resistant weeds[40]. These combined attributes make HPPD inhibitors an important tool in diversified weed management programs.

Palmer amaranth, an invasive annual broadleaf weed native to the southwestern United States, has emerged as a significant threat to agricultural crops and ecosystems worldwide[42, 43]. Its adaptability, aggressive growth, and resistance to herbicides have made it a troublesome problem in various regions[43, 44]. With distinct male and female plants, Palmer amaranth efficiently cross-pollinates, facilitating the exchange of genetic material and the emergence of novel traits[44]. These biologic characteristics of this species led to resistance evolution to multiple herbicide modes of action, including ALS-, EPSPS-, synthetic auxins, PPO-, PSII-, HPPD-, VLCFA-, and tubulin-inhibitors[7]. To enhance our understanding of the genetic factors underlying resistance and other crucial traits in Palmer amaranth, further investigation in resistant biotypes is necessary.

The present dissertation is structured into three chapters, each focusing on different aspects of metabolic herbicide resistance. Chapter 1 provides a comprehensive review of metabolic herbicide resistance, which is considered the most significant mechanism among NTSR mechanisms in weeds[8]. This chapter encompasses key milestones in research on metabolic resistance in weeds, as well as the identification of crucial genes associated with this resistance mechanism. Chapter 2 presents a research study conducted on a specific population of HPPD-resistant Palmer amaranth, known as NER. This population has been characterized by its resistance to tembotrione through enhanced metabolism[36]. The chapter encompasses RNA-seq analysis, utilization of heterologous systems and other experiments to identify and validate the

genes responsible for the resistant trait. Additionally, the chapter explores the *cis*- and *trans*-elements that may be involved in the resistance mechanism. The third and final chapter focuses on elucidating the role of P450 reductase in relation to metabolic resistance in weeds. The objective of this chapter is to gain a better understanding of the involvement of this protein in metabolic resistance and explore its potential as a target for developing new management tools for weed control.

REFERENCES

1. FAO. Global agriculture towards 2050. In *How to feed the world 2050*: Rome (2009).
2. Oerke EC, Crop losses to pests. *Journal of Agricultural Science* **144**: 31-43 (2006).
3. Gallandt ER and Weiner J. Crop–weed competition. In *Encyclopedia of Life Sciences*, pp. 1-9.
4. van Heemst HDJ, The influence of weed competition on crop yield. *Agricultural Systems* **18**: 81-93 (1985).
5. Peterson MA, Collavo A, Ovejero R, Shivrain V and Walsh MJ, The challenge of herbicide resistance around the world: a current summary. *Pest Manag Sci* **74**: 2246-2259 (2018).
6. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A and Dayan FE, Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry* **295**: 10307–10330 (2020).
7. Heap I, The international survey of herbicide resistant weeds. http://www.weedscience.com [accessed Jan 10, 2023 2022].
8. Rigon CAG, Gaines TA, Küpper A and Dayan FE, Metabolism-based herbicide resistance, the major threat among the non-target site resistance mechanisms. *Outlooks on Pest Manag* **31**: 162-168 (2020).
9. Powles SB and Yu Q, Evolution in Action: Plants Resistant to Herbicides. *Annu Rev Plant Biol* **61**: 317-347 (2010).
10. Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE and Westra P, Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences* **107**: 1029-1034 (2010).
11. Laforest M, Soufiane B, Simard MJ, Obeid K, Page E and Nurse RE, Acetyl-CoA carboxylase overexpression in herbicide-resistant large crabgrass (*Digitaria sanguinalis*). *Pest Manag Sci* **73**: 2227-2235 (2017).
12. Yu Q, Abdallah I, Han HP, Owen M and Powles S, Distinct non-target site mechanisms endow resistance to glyphosate, ACCase and ALS-inhibiting herbicides in multiple herbicide-resistant *Lolium rigidum*. *Planta* **230**: 713-723 (2009).
13. Moretti ML and Hanson BD, Reduced translocation is involved in resistance to glyphosate and paraquat in *Conyza bonariensis* and *Conyza canadensis* from California. *Weed Res* **57**: 25-34 (2017).
14. Ge X, d'Avignon DA, Ackerman JJH and Sammons RD, Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. *Pest Manag Sci* **66**: 345-348 (2010).
15. Powles SB and Yu Q. Evolution in action: plants resistant to herbicides. In *Annu Rev Plant Biol*, ed. by Merchant S, Briggs WR and Ort D. Annual Reviews: Palo Alto, pp. 317-347 (2010).
16. Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S and Werck-Reichhart D, Cytochromes p450. *Arabidopsis Book* **9**: e0144 (2011).
17. Edwards R and Dixon DP, Plant glutathione transferases. *Methods Enzymol* **401**: 169-186 (2005).
18. Pan L, Yu Q, Wang J, Han H, Mao L, Nyporko A, Maguza A, Fan L, Bai L and Powles S, An ABCC-type transporter endowing glyphosate resistance in plants. *Proceedings of the National Academy of Sciences* **118**: e2100136118 (2021).
19. Brunharo C and Hanson BD, Vacuolar sequestration of paraquat is involved in the resistance mechanism in *Lolium perenne* L. spp. *multiflorum*. *Front Plant Sci* **8**: 9 (2017).
20. Goldberg-Cavalleri A, Onkokesung N, Franco-Ortega S and Edwards R, ABC transporters linked to multiple herbicide resistance in blackgrass (*Alopecurus myosuroides*). *Front Plant Sci* **14**: 1082761 (2023).
21. Iwakami S, Endo M, Saika H, Okuno J, Nakamura N, Yokoyama M, Watanabe H, Toki S, Uchino A and Inamura T, Cytochrome P450 *CYP81A12* and *CYP81A21* are associated with resistance to two acetolactate synthase inhibitors in *Echinochloa phyllopogon*. *Plant Physiol* **165**: 618-629 (2014).

22. Dimaano NG, Yamaguchi T, Fukunishi K, Tominaga T and Iwakami S, Functional characterization of cytochrome P450 *CYP81A* subfamily to disclose the pattern of cross-resistance in *Echinochloa phyllopogon*. *Plant MolBiol* **102**: 403-416 (2020).
23. Zheng T, Yu X, Sun Y, Zhang Q, Zhang X, Tang M, Lin C and Shen Z, Expression of a cytochrome P450 gene from Bermuda grass *Cynodon dactylon* in soybean confers tolerance to multiple herbicides. *Plants* **11**: 949 (2022).
24. Han H, Yu Q, Beffa R, González S, Maiwald F, Wang J and Powles SB, Cytochrome P450 *CYP81A10v7* in *Lolium rigidum* confers metabolic resistance to herbicides across at least five modes of action. *The Plant Journal* **105**: 79-92 (2021).
25. Shen J, Yang Q, Hao L, Zhang L, Li X and Zheng M, The metabolism of a novel cytochrome P450 (*CYP77B34*) in tribenuron-methyl-resistant *Descurainia sophia* L. to herbicides with different mode of actions. *Int J Mol Sci* **23**: 5812 (2022).
26. Lu H, Liu Y, Li M, Han H, Zhou F, Nyporko A, Yu Q, Qiang S and Powles S, Multiple metabolic enzymes can be involved in cross-resistance to 4-hydroxyphenylpyruvate-dioxygenase-inhibiting herbicides in wild radish. *J Agric Food Chem* 2023).
27. Jensen K and Møller BL, Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* **71**: 132-141 (2010).
28. Rana S, Lattoo SK, Dhar N, Razdan S, Bhat WW, Dhar RS and Vishwakarma R, NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from Withania somnifera (L.) dunal. *PLoS One* **8**: e57068 (2013).
29. Yamamura Y and Mabuchi A, Functional characterization of NADPH-cytochrome P450 reductase and cinnamic acid 4-hydroxylase encoding genes from *Scoparia dulcis* L. *Botanical Studies* **61**: 6 (2020).
30. Urban P, Mignotte C, Kazmaier M, Delorme F and Pompon D, Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 *CYP73A5*. *Journal of Biological Chemistry* **272**: 19176-19186 (1997).
31. Mizutani M and Ohta D, Two isoforms of NADPH:cytochrome P450 reductase in Arabidopsis thaliana. Gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol* **116**: 357-367 (1998).
32. Rana S, Lattoo SK, Dhar N, Razdan S, Bhat WW, Dhar RS and Vishwakarma R, NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from *Withania somnifera* (L.) Dunal. *PLoS One* **8**: 16 (2013).
33. Parage C, Foureau E, Kellner F, Burlat V, Mahroug S, Lanoue A, Dugé de Bernonville T, Londono MA, Carqueijeiro I, Oudin A, Besseau S, Papon N, Glévarec G, Atehortùa L, Giglioli-Guivarc'h N, St-Pierre B, Clastre M, O'Connor SE and Courdavault V, Class II cytochrome P450 reductase governs the biosynthesis of alkaloids. *Plant Physiol* **172**: 1563-1577 (2016).
34. Jhala AJ, Kumar V, Yadav R, Jha P, Jugulam M, Williams MM, Hausman NE, Dayan FE, Burton PM, Dale RP and Norsworthy JK, 4-Hydroxyphenylpyruvate dioxygenase (HPPD)-inhibiting herbicides: past, present, and future. *Weed Technol* **37**: 1-14 (2023).
35. Lu H, Yu Q, Han H, Owen MJ and Powles SB, Evolution of resistance to HPPD-inhibiting herbicides in a wild radish population via enhanced herbicide metabolism. *Pest Manag Sci* **76**: 1929-1937 (2019).
36. Küpper A, Peter F, Zollner P, Lorentz L, Tranel PJ, Beffa R and Gaines TA, Tembotrione detoxification in 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor-resistant Palmer amaranth (*Amaranthus palmeri* S. Wats.). *Pest Manag Sci* **74**: 2325-2334 (2018).
37. McMullan PM and Green JM, Identification of a Tall waterhemp (*Amaranthus tuberculatus*) biotype resistant to HPPD-inhibiting herbicides, atrazine, and thifensulfuron in Iowa. *Weed Technol* **25**: 514-518 (2011).
38. Hawkes TR, Langford MP, Viner R, Blain RE, Callaghan FM, Mackay EA, Hogg BV, Singh S and Dale RP, Characterization of 4-hydroxyphenylpyruvate dioxygenases, inhibition by

herbicides and engineering for herbicide tolerance in crops. *Pest Biochem Physiol* **156**: 9-28 (2019).
39. Santucci A, Bernardini G, Braconi D, Petricci E and Manetti F, 4-Hydroxyphenylpyruvate dioxygenase and its inhibition in plants and animals: small molecules as herbicides and agents for the treatment of human inherited diseases. *Journal of Medicinal Chemistry* **60**: 4101-4125 (2017).
40. Almsick Av, New HPPD-Inhibitors – A proven mode of action as a new hope to solve current weed problems. *Outlooks on Pest Management* **20**: 27-30 (2009).
41. Ndikuryayo F, Moosavi B, Yang W-C and Yang G-F, 4-Hydroxyphenylpyruvate dioxygenase inhibitors: from chemical biology to agrochemicals. *J Agric Food Chem* **65**: 8523-8537 (2017).
42. Ward SM, Webster TM and Steckel LE, Palmer Amaranth (*Amaranthus palmeri*): A Review. *Weed Technol* **27**: 12-27 (2013).
43. Yanniccari M, Gaines T, Scursoni J, Prado Rd and Vila-Aiub M, Global patterns of herbicide resistance evolution in *Amaranthus* spp .: an analysis comparing species, cropping regions and herbicides. **40**: - (2022).
44. Roberts J and Florentine S, A review of the biology, distribution patterns and management of the invasive species *Amaranthus palmeri* S. Watson (Palmer amaranth): Current and future management challenges. *Weed Res* **62**: 113-122 (2022).

2. CHAPTER I - HERBICIDE METABOLISM IN WEEDS: WHAT WE'VE LEARNT

INTRODUCTION

Herbicide resistance in weeds poses a significant challenge to modern agriculture worldwide, compromising effective weed control strategies[1]. Herbicide resistance can arise through two main mechanisms: target-site (TSR) and non-target-site resistance (NTSR)[2]. TS resistance involves mutations in the herbicide's target site, reducing its binding affinity and rendering the herbicide ineffective[2]. NTSR, on the other hand, encompasses a diverse range of mechanisms that facilitate the detoxification, sequestration, or altered translocation of herbicides within weed plants, ultimately reducing their efficacy. Although both TSR and NTSR contribute to herbicide resistance in weeds, the metabolic mechanism stands out as the major and more complex resistance mechanism, owing to its extensive capacity to metabolize a wide range of herbicides within weed species[3, 4].

Metabolic resistance in weeds is a process involving the enzymatic detoxification and metabolism of xenobiotics, including herbicides[4, 5]. This process is divided into three key phases: Phase I - activation, phase II - conjugation, and phase III – sequestration (Figure 2-1). During the activation phase, xenobiotics, such as herbicides, undergo chemical modifications to generate more reactive forms. In this phase, cytochrome P450 monooxygenases (P450s) play a central role in metabolizing herbicides by adding a small functional group on the herbicide molecule. P450 enzymes catalyze various reactions, including hydroxylation, dealkylation, deamination and others, which modify the chemical structure of herbicides and reduce their toxicity[2]. In the phase II, other enzymes such as glutathione *S*-transferases (GST) and esterases, contribute to herbicide detoxification through conjugation and hydrolysis reactions, respectively[6]. In the phase III, modified herbicides or their metabolites are sequestered and compartmentalized within

specific cellular compartments, mainly by ABC transporters, preventing their interaction with essential plant targets[2, 3].

In the last decade, there has been a growing recognition of the increasing importance of metabolic resistance mechanisms in weeds. The emergence and spread of herbicide-resistant weed populations have posed significant challenges to sustainable agriculture and weed management practices. While TSR has long been studied and documented, the significance of metabolic resistance mechanisms has gained prominence due to their ability to confer broad-spectrum resistance to multiple herbicides. The complexity and diversity of metabolic pathways involved in herbicide metabolism highlight the adaptability and evolutionary potential of weeds. Understanding and addressing the metabolic resistance mechanisms in weeds have become crucial for developing effective strategies to combat herbicide resistance and enable global crop production. Our current understanding of metabolic resistance in weeds has significantly advanced in recent decades, thanks to various scientific breakthroughs discussed in the next section.

MILESTONES IN METABOLIC RESISTANCE IN WEEDS

In 1957, the first herbicide resistant weed was documented, in wild carrot (*Daucus carota*) with resistance to 2,4-D[7] (Figure 2-2). Subsequently, additional cases of resistance emerged; however, the underlying resistance mechanisms remained elusive. The introduction and widespread use of acetyl-CoA carboxylase (ACCase) and acetolactate synthase (ALS) inhibitors in the early 1970s and 1980s, respectively, resulted in a substantial increase in reported cases of herbicide resistance in weeds. Besides the growing number of reported cases, many of them were not extensively investigated.

The continued use of selective herbicides in early 1980s, especially ALS-inhibitors in cereals, selected weeds resistant to herbicides. The mechanisms of selectivity in crops were well known to be enhanced metabolism[8]. However, cases of weeds resistant to different herbicides started to arise. The term cross-resistance was not utilized in weeds but well documented and understood in insects[9]. At that time, the first cases of weeds with cross-resistance start to evolve and scientist were worried that resistant biotypes were cross-resistant to different chemical families and in some cases cross-resistant to herbicide never used before. More than 20 years after the first weed resistance report, one of the first papers published assuming metabolic resistance was evolved in ryegrass (*Lolium rigidum*) biotype from Australia collected in 1980 (Figure 2-2). This biotype was resistanct to diclofop, fluazifop, haloxyfop, quizalofop, sethoxydim, tralkoxydim (ACCase inhibitors) and cross-resistant to chlorsulfuron (ALS inhibitor) and trifluralin (Microtubule inhibitor)[10, 11]. A similar resistance mechanism was identified in a blackgrass (*Alopecurus myosuroides*) biotype at Peldon, Essex, United Kingdom in 1984. This biotype was resistant to chlortoluron, and interestingly it was cross-resistant to pendimethalin and diclofop-methyl[12]. The levels of resistance for these cases were lower if compared to the ones found in triazine resistance plants[13], but even though the problem of cross-resistance was a big concern in that time due to cross-resistance to herbicides that were never even applied in the field before[12].

At that time, the knowledge about metabolic resistance was limited. The use of P450 inhibitors like 1-aminobenzotriazole (ABT), piperonylbutoxide (PBO), and malathion were used to start the metabolic resistance studies. In both *Lolium* and *Alopecurus*, the herbicides had synergistic effects with P450 inhibitors on controlling the resistant weeds, showing the involvement of metabolism in the resistance mechanism[12, 14]. These cases were the first cases of

involvement of metabolic resistance in weeds. However, during the middle 1980s and early 1990s the use of ^{14}C herbicides and identification of the metabolites produced by resistant weeds helped to confirm the metabolic resistance and to infer about which metabolic enzymes were involved. Resistance to chlortoluron and isoproturon in blackgrass population from Peldon was attributed to metabolism due to a faster degradation of C^{14} herbicides in the leaves to different metabolites[12]. *Lolium rigidum* were identified to be resistant to simazine (Photosystem II inhibitor, PSII) by a faster rate of metabolism of simazine to N-de-ethyl derivates[15], probably done by P450 enzymes. Population of *L. rigidum* and *A. myosuroides* resistant to chlorsulfuron and chlortoluron, respectively, had the same detoxification pathway as wheat, by hydroxylation of the phenyl-ring followed by conjugation to glucose in *L. rigidum* [16], or by N-demethylation in *A. myosuroides* [17] inferring the participation of P450 and GSTs enzymes. Populations of *Abutilon theophrasti* resistant to atrazine were identified to conjugate atrazine in a faster rate than susceptible plants, showing the involvement of GST enzymes in the metabolic resistance mechanism [18].

With the passage of time and repeated herbicide use, new cases of metabolic resistance have been identified in different species. Metabolic resistance in *Amaranthus hybridus* was reported in 1999. Resistant plants metabolized ^{14}C-chlorimuron faster than susceptible biotypes [19]. The first report of enhanced metabolism in *Lolium multiflorum* was reported in 2001. Resistant plants from UK rapidly detoxified diclofop by glycosyl ester formation and was cross-resistant to tralkoxydim and cycloxydim [20]. In early 2000, a population of *Avena* spp, evolved resistance to fenoxaprop and cross resistant to tralkoxydim [21]. *Bromus tectorum* was reported resistant to propoxycarbazone (ALS inhibitor) by rapid rate of metabolism compared to sensitive plants and probably P450 enzymes were involved in the mechanism [22].

Questions about the evolution of metabolic resistance in weeds were arising. It was suggested that high pesticide rates would select major genes, as target-site resistance had been selected. However, applications of low herbicide rates were suggested to evolve polygenic herbicide resistance [23]. One notable experiment on metabolic resistance was a study using recurrent selection with herbicide low rates in *L. rigidum* (Figure 2-2). This experiment was performed in middle 2000s and was conducted with susceptible plants. After only three cycles of plant selection using low doses of diclofop, the progenies had high level of resistance by enhanced metabolism. More than that, the progenies evolved cross-resistance [24]. These results were a milestone on this subject and showed the rapid evolution of this type of mechanism of resistance and the consequence of using the same herbicide mode of action for consecutive years. It proved that low rates could favor the selection of minor alleles and hence stack up in the progenies and evolve resistance due to a polygenic trait.

Early studies on herbicide metabolism in weeds primarily relied on in vivo P450 inhibitors, lacking evidence of detoxification enzymes. In the 1990s, microsome isolation from crops facilitated in vitro investigations of P450-mediated metabolism with various herbicides[25]. It took around 15 more years before microsomes isolation from herbicide resistant weeds was performed. In 2005, isolation of microsome from resistant *Echinochloa phyllopogon* showed that microsomes had higher P450 activity when incubated with bispyribac, fenoxaprop and thiobencarb, which may correspond with the hydroxylation of the herbicide molecules [26]. Heterologous expression in yeast was another method used to understand P450 function and substrate affinity (Figure 2-2). Numerous P450 genes from crops were randomly isolated, expressed in yeast and subjected to incubation with herbicides[25]; however, due to the limited

availability of isolated P450 enzymes from weeds at that time, this method was not extensively employed for weed-related studies.

Around 2010, through the isolation and identification of highly induced genes in the resistant population of *Echinochloa phyllopogon*, two significant genes, *CYP81A12* and *CYP81A21*, were discovered as the first weed genes capable of conferring herbicide metabolism (Figure 2-2). By transforming these genes into *Arabidopsis thaliana* plants, resistance to multiple herbicides was observed, including bensulfuron-methyl through O-demethylation and penoxsulam[27]. Same results were observed when these genes were heterologous expressed in yeast[27]. Recently, recombinant expression of these *CYPs* from 81A family in *E. coli* has demonstrated their ability to metabolize a wide range of herbicides, including those targeting ALS, ACCase, phytoene desaturase (PDS), 4-Hydroxyphenylpyruvate dioxygenase (HPPD), Protoporphyrinogen oxidase (PPO) and PSII inhibitors[28].

At the same time, nearly 2014, whole transcriptome analysis started to be performed to compare the transcripts levels between resistant and susceptible weeds aiming to identify metabolic candidate genes (Figure 2-2). The first study was performed in *L. rigidum* resistant to diclofop[29]. Subsequently, numerous studies have utilized RNA-seq analysis to identify candidate genes responsible for metabolic resistance in various weed species, including *Lolium* spp. [30], *Polypogon fugax* [31], *Alopecurus aequalis* [32], *Echinochloa crus-galli* [33] and others.

The identification and the advanced biotechnologies make it easier, but not less challenging, to validate metabolic candidate genes found in RNA-seq analysis using heterologous expression systems. Jumping to recent years, rice and Arabidopsis are the most common plants used for transformation. As cited before, Arabidopsis carrying *CYP81A12* and *CYP81A21* from *E. phyllopogon* were the first successful validation of metabolic genes from

weeds. One case that was important in the history of metabolic resistance in weeds was found in an *Echinochloa colona* biotype resistant to glyphosate. Glyphosate is the most commonly used herbicide in the world and was not known to be metabolized by plants. This *E. colona* biotype evolved resistance to glyphosate and TSR was discarded. By RNA-seq analysis, the authors identified an aldo-keto-reductase (*AKR*) gene that was upregulated in resistant plants. Rice plants were transformed with this *AKR* from *E. colona* and became resistant to glyphosate. The mechanism of resistance is based on oxidation reaction of glyphosate to AMPA performed by *AKR*. This result was a milestone because glyphosate is basically never detoxified by plants [34]. In the same population of *E. colona*, an ABC transporter gene (*EcABCC8*) was identified in the RNA-seq analysis and transformed rice, maize, and soybean overexpressing *EcABCC8* became glyphosate resistant. The authors used CRISPR/Cas9 to knockout the OsABCC8 in rice and the plants became resistant to glyphosate. The mechanism of resistance is based on movement of glyphosate molecules to extracellular space of the cell (apoplast)[35].

With advancements in technology, the identification of the underlying causal genes responsible for metabolic resistance in most weed species has become increasingly feasible. The availability of genome sequences for weed species, with the expectation of more through initiatives like the International Weed Genomics Consortium (https://www.weedgenomics.org/), has opened new possibilities for research. Previously unattainable studies such as QTL mapping can now be conducted in weeds. The first QTL mapping study was performed in *Amaranthus tuberculatus* with metabolic resistance to HPPD-inhibitors[36]. Such studies contribute to understanding the regulation of metabolic genes, representing the next step in researching metabolic mechanisms.

Since 1980, research has been focused on the identification of metabolic resistance in grasses. Recently, metabolic resistance has been reported in various important dicot weed species. Metabolic resistance was reported in *A. tuberculatus* in 2017 [37], *A. palmeri* [38] in 2018, and *Raphanus raphanistrum* in 2020 [39], all of which exhibited resistant to HPPD inhibitors. Additionally, in 2020, *A. tuberculatus* was identified as resistant to 2,4-D through metabolic resistance[40]. Early efforts involved the use of P450 inhibitors, ^{14}C herbicides, and the analysis of metabolites. However, we have now entered an era of gene identification, validation and regulation[2, 3, 41]. Recently, researchers have successfully identified and validated numerous genes as potential contributors to metabolic resistance (detailed in the next section). To gain further insights into the underlying mechanisms, protein modeling techniques were employed, particularly for P450, to better understand substrate affinity and identify potential substrate binding sites. Looking ahead, we anticipate that in the coming years, advancements in our understanding will allow us to identify specific genes and predict their broad substrate recognition solely based on their DNA sequences. However, it is important to note that our knowledge of the genetic regulation of these resistance-related genes in plants is still incomplete. Therefore, we believe that future studies should prioritize investigating the genetic regulation of these genes in order to enhance our understanding of metabolic resistance. Overall, research in metabolism-based resistance has made significant strides in unraveling the complexities of metabolic resistance in both grasses and dicots. The future is optimistic about the prospects of gene identification and substrate recognition and remain committed to furthering our understanding of the genetic regulation underlying metabolic resistance in plants.

GENES INVOLVED IN METABOLIC HERBICIDE RESISTANCE

In the field of herbicide resistance, significant progress has been made in the identification and functional validation of numerous genes associated with herbicide metabolism in both resistant weeds and crops. This section focuses on the key genes that have been discovered in weeds and crops, demonstrating their ability to metabolize different herbicides. The identification and validation of these important genes in both weeds and crops represent significant advancements in our knowledge of metabolic resistance and provide a foundation for further research in this field. By understanding the specific genes families involved in herbicide metabolism, researchers can gain insights into the mechanisms underlying resistance and potentially develop strategies to overcome it.

The cytochrome P450 (P450) gene family is widely recognized as the predominant detoxification gene family in plants[42]. Arabidopsis and rice have 244 and 326 cytochrome P450[42, 43], respectively, which exhibit diverse functions and play essential roles in various metabolic pathways, including the metabolism of herbicide. In weeds the number of P450 genes are overrepresented in some species. For example, *Echinochloa crus-galli* and *Echinochloa colona* have a total of 867 and 694 P450s, respectively, which is significantly higher when compared with other species[44]. The extensive number of P450 genes in weeds highlights their metabolic versatility and its potential to adapt to various environmental challenges.

Cytochrome P450 genes from various families have been identified and validated in both crops and weeds (Figure 2-3). In crops like rice, *OsCYP72A31* and *OsCYP76C6* were found to metabolize ALS-inhibitors such as bispyribac and bensulfuron[45], while *OsCYP81A6* exhibited a broad spectrum of activity, metabolizing several ALS-inhibitors and clomazone[46, 47]. In soybean, a mutation in the *GmCYP81E22* gene was identified as the cause of high sensitivity to bentazon

in specific cultivars[48] and *GmCYP71A10*, when expressed in yeast and transformed into tobacco plants, demonstrated the ability to metabolize Chlortoluron [49]. Among corn genes, *ZmCYP81A2* was found to hydroxylate bentazon at multiple positions, and *ZmCYP81A9* exhibited a wide range of activity, metabolizing HPPD, ALS, and PSII-inhibitors[50]. In weeds, CYPs from multiple herbicide resistant *Echinochloa phyllopogon* (e.g., *EpCYP81A12*, *EpCYP81A21* and *EpCYP81A24*) metabolized a wide range of substrate metabolism, including ACCase, ALS, PSII, DXS, PDS, and HPPD-inhibitors[27, 28]. *EpCYP81A14* and *EpCYP81A18* shared high similarity with the previous enzymes but exhibited different substrate specificity, metabolizing only ALS-inhibitors[28]. Similarly, *Cynodon dactylon* genes *CdCYP81A69* and *CdCYP81A70*, from the same family, when transformed into soybean, conferred resistance to ALS inhibitors and 2,4-D[51]. Another gene with a broad spectrum was found in multiple herbicide-resistant annual ryegrass, where transgenic rice overexpressing *Lolium LrCYP81A10v7* exhibited high resistance to ACCase, ALS, and moderate resistance to PSII, HPPD, and tubulin inhibitors [52]. In other hand, the genes from *Beckmannia syzigachne* including *BsCYP81Q32*, *BsCYP704A177* and *BsCYP99A44* when transformed in Arabidopsis plants confer herbicide resistance specific to ALS inhibitor herbicides[53, 54]. In dicot weed species, fewer P450 genes have been identified. For instance, *DsCYP77B34* from *Descurainia sophia*, a broadleaf weed in Asian wheat fields, rendered transformed Arabidopsis resistant to ALS-, PPO-, very long chain fatty acid (VLCFA-), and PSII-inhibitors[55]. Arabidopsis expressing *Raphanus raphanistrum RrCYP704C1* or Rr*CYP709B1* demonstrated resistance to HPPD inhibitors such as mesotrione (*RrCYP704C1*), tembotrione (*RrCYP709B1*), and isoxaflutole (*RrCYP709B1*)[56]. The evolution of multiple genes in grass weed species can be attributed to prolonged exposure to herbicide modes of action, particularly ACCase and ALS-inhibitors, over the years. It is evident that the P450 family 81

plays a pivotal role in grasses, while other families are found in both grasses and dicots (Figure 2-3).

Fewer genes have been identified and validated for phase II and III of herbicide metabolism, with significant research focused on corn, soybean, and wheat (Figure 2-3). In these crops, efforts have been made to isolate and study *GSTs*. Expression of various *Phi* and *Tau* *ZmGSTUs*, *Phi TaGSTs*, and *Tau* class *GmGSTUs* in *E. coli* demonstrated some degree of conjugation to ALS-, PPO-, and VLCFA-inhibitors[57-59]. Among weeds, only *AmGSTF1* from *Alopecurus myosuroides* has been identified and validated. This black-grass population exhibited multiple herbicide resistance, and transgenic Arabidopsis overexpressing *AmGSTF1* displayed resistance to Chlortoluron (ALS), Alachlor (VLCFA), and atrazine (PSII), the same herbicides to which the resistant black-grass plants were resistant[60, 61]. Knocking out *AmGSTF1* through virus-mediated transient expression in resistant black-grass plants restored sensitivity to fenoxaprop (ACCase), confirming the essential role of *AmGSTF1* in herbicide resistance [61]. The genes associated with phase III, such as ABC transporters, which sequesters herbicides into the vacuole, have received limited attention as mechanisms of resistance. In glyphosate resistant *Echinochloa colona*, it was found that *ABCC8* was constitutively expressed at higher levels in resistant plants. Transgenic rice expressing this gene exhibited resistance to glyphosate by extruding the herbicide from the cytoplasm to the apoplast[35]. Similarly, three ABC transporters (*AmABCC1*, *AmABCC2*, *AmABCC3*) was recently identified in multiple-resistant populations of *Alopecurus myosuroides*. Transgenic yeast carrying *AmABCC1* or *AmABBC2* had enhanced tolerance to mesosulfuron-methyl (ALS), highlighting the role of these proteins in sequestering the herbicide into the vacuole in resistant blackgrass[62].

Ongoing research advancements are expected to uncover more genes, especially in dicots, benefiting from advancements in technology and genome availability. Comprehending their functions and correlation will aid in combating herbicide resistance evolution. Armed with this understanding, we can devise focused approaches to efficiently manage weeds and impede resistance development. Through the utilization of advanced tools and scientific methodologies, we are steadily advancing towards sustainable agriculture and improved weed control.

FUNCTIONAL GENOMICS

Functional genomics is playing a crucial role in agronomic weed science, particularly in understanding evolution of different resistance mechanisms. With the advancements in high-throughput sequencing technologies, scientists can now obtain complete genomes of weed species more efficiently. These new genomic resources will provide a wealth of information about the genetic variations and functional elements in weed populations, enabling researchers to identify key genes involved in metabolic resistance and develop targeted management approaches.

Scientists employ various techniques to study functional genomics in the context of weed science. One commonly used approach is transcriptomics, which involves analyzing the expression patterns of genes in different weed populations, including both resistant and susceptible biotypes. This helps identify genes that are upregulated or downregulated in resistant weeds, providing insights into the metabolic pathways and processes involved in herbicide resistance[29, 32]. Another technique is proteomics, which involves studying the complete set of proteins produced by weed species[63]. Proteomics can shed light on the proteins involved in herbicide metabolism and target-site interactions, contributing to a better understanding of resistance mechanisms. Additionally, metabolomics, the study of small molecule metabolites,

allows researchers to explore the metabolic changes associated with resistance[64]. By applying these techniques, scientists can gather comprehensive data on the functional elements of weed genomes and their interactions with herbicides.

Gene validation techniques such as Arabidopsis or tobacco transformation, virus-induced gene silencing (VIGS), and mutants database are valuable tools for determining gene functions involved in the evolution of herbicide resistance in weeds as it was described in the section 2. Furthermore, recent advancements in gene editing technologies such as RNA interference (RNAi) and CRISPR have revolutionized our ability to understand and address herbicide resistance in weeds[65, 66]. RNAi allows for targeted gene knockdown, enabling researchers to assess the impact of specific genes on resistance phenotypes[66]. CRISPR, on the other hand, enables precise gene editing, facilitating the study of gene functions and the discovery of new targets for sustainable weed management. These technologies hold great promise for both unraveling the mechanisms behind metabolic resistance evolution and developing innovative strategies for weed control.

CONCLUSIONS AND PERSPECTIVES

Herbicide resistance in weeds poses a global challenge to agriculture, necessitating a deep understanding of resistance mechanisms, especially metabolic resistance, for effective weed management. The future holds promising prospects for gene identification, substrate recognition, and the prediction of broad substrate recognition based on DNA sequences. Advancements in technology and the availability of genomic databases are expected to expedite gene discovery, particularly in dicot weed species. Investigating the genetic regulation of resistance-related genes should be a priority to enhance our comprehension of metabolic resistance.

Continuous advancements in metabolic research provide valuable insights into the complexities of metabolic resistance in weeds. Combined with an improved understanding of genetic regulation, this knowledge will aid in combating the evolution of herbicide resistance and developing targeted strategies for effective weed management. By leveraging advanced tools and scientific methodologies, we are making strides toward achieving sustainable agriculture and enhanced weed control practices.

In addition, the public-private partnership model, exemplified by the International Weed Genomics Consortium (IWGC), plays a crucial role in advancing weed science. Through efforts such as hosting comprehensive databases, integrating new datasets specific for metabolic resistance, providing training, and facilitating research coordination, ii promotes the sharing of advancements in metabolic resistance research across the scientific community. Utilizing this collaborative approach will further enhance our understanding of metabolic resistance and contribute to improved weed management strategies.

By embracing these perspectives and leveraging the power of interdisciplinary collaboration, we can effectively address the challenges posed by herbicide resistance in weeds, ensuring sustainable agriculture and enhanced weed control practices for the future.

FIGURES

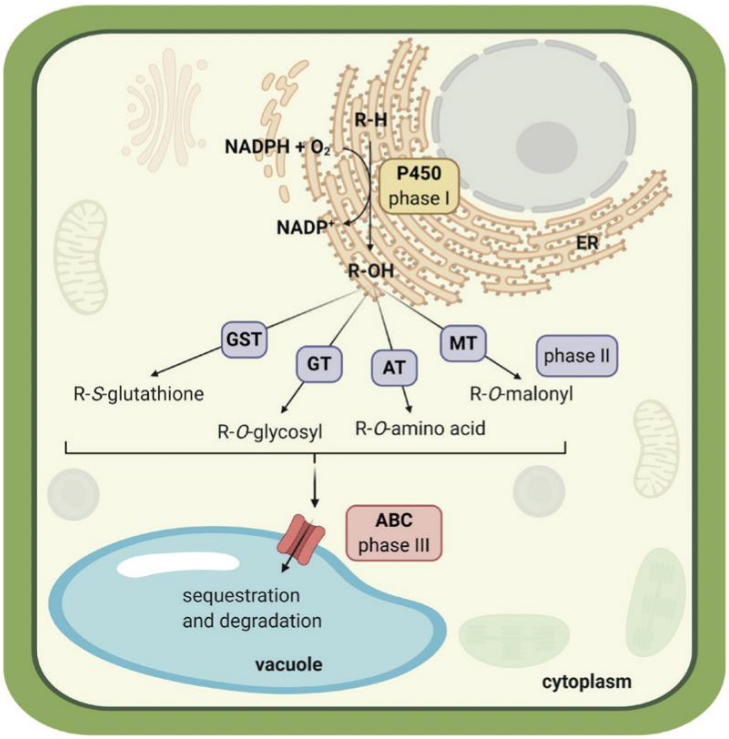


Figure 2-1. Phases of herbicide metabolic process, in which herbicides can undergo when inside plant cells. Note: Figure from "Metabolism-based herbicide resistance, the major threat among the non-target site resistance mechanisms", by C. A. G. Rigon, T. A. Gaines, A. Küpper and F. E. Dayan, Outlooks on Pest Manag. 2020. Vol. 31 Issue 4 Pages 162-168. DOI: 10.1564/v31_aug_04.

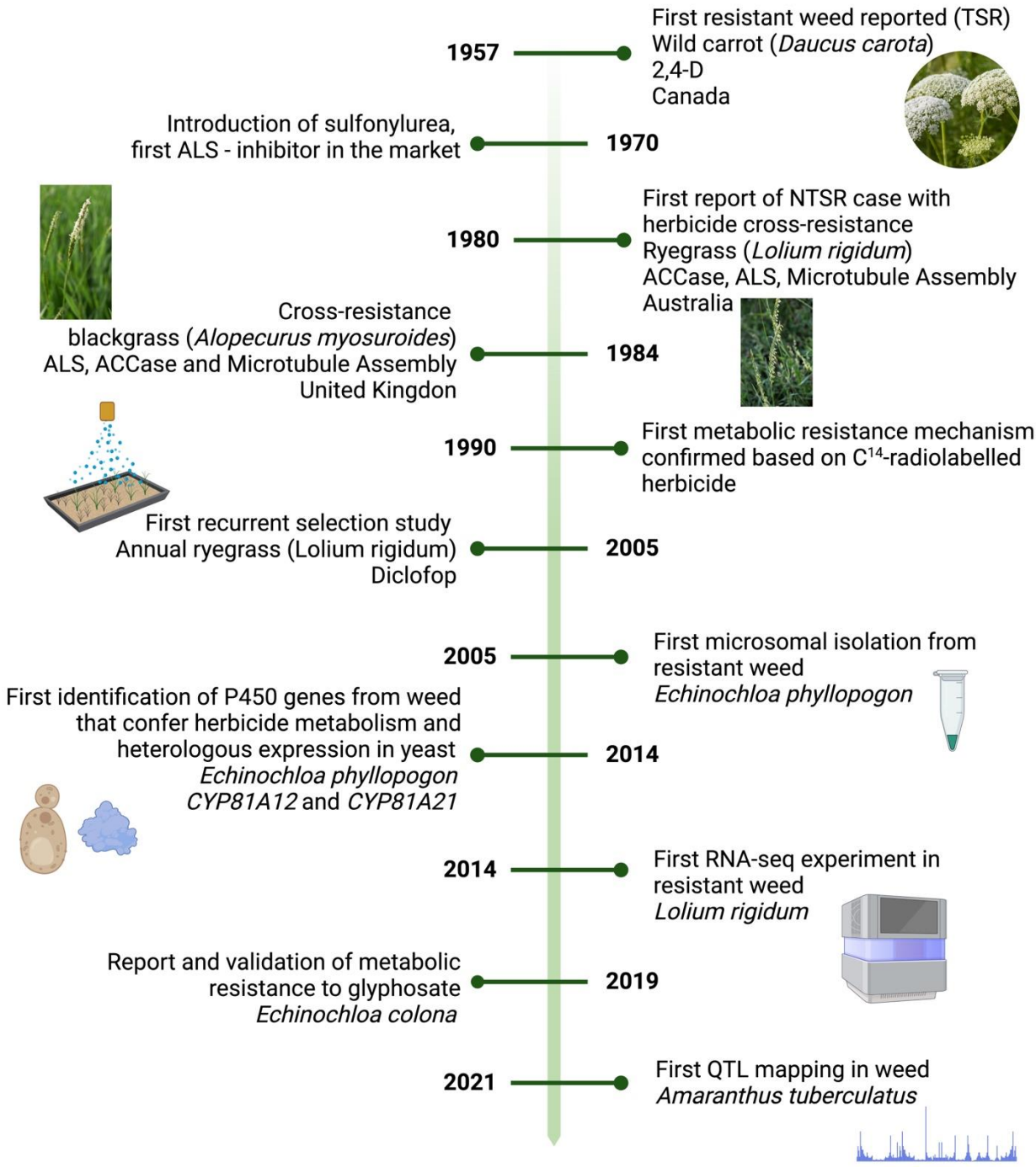


Figure 2-2. Timeline of key milestones in understanding metabolic resistance mechanisms in weeds over the past decades.

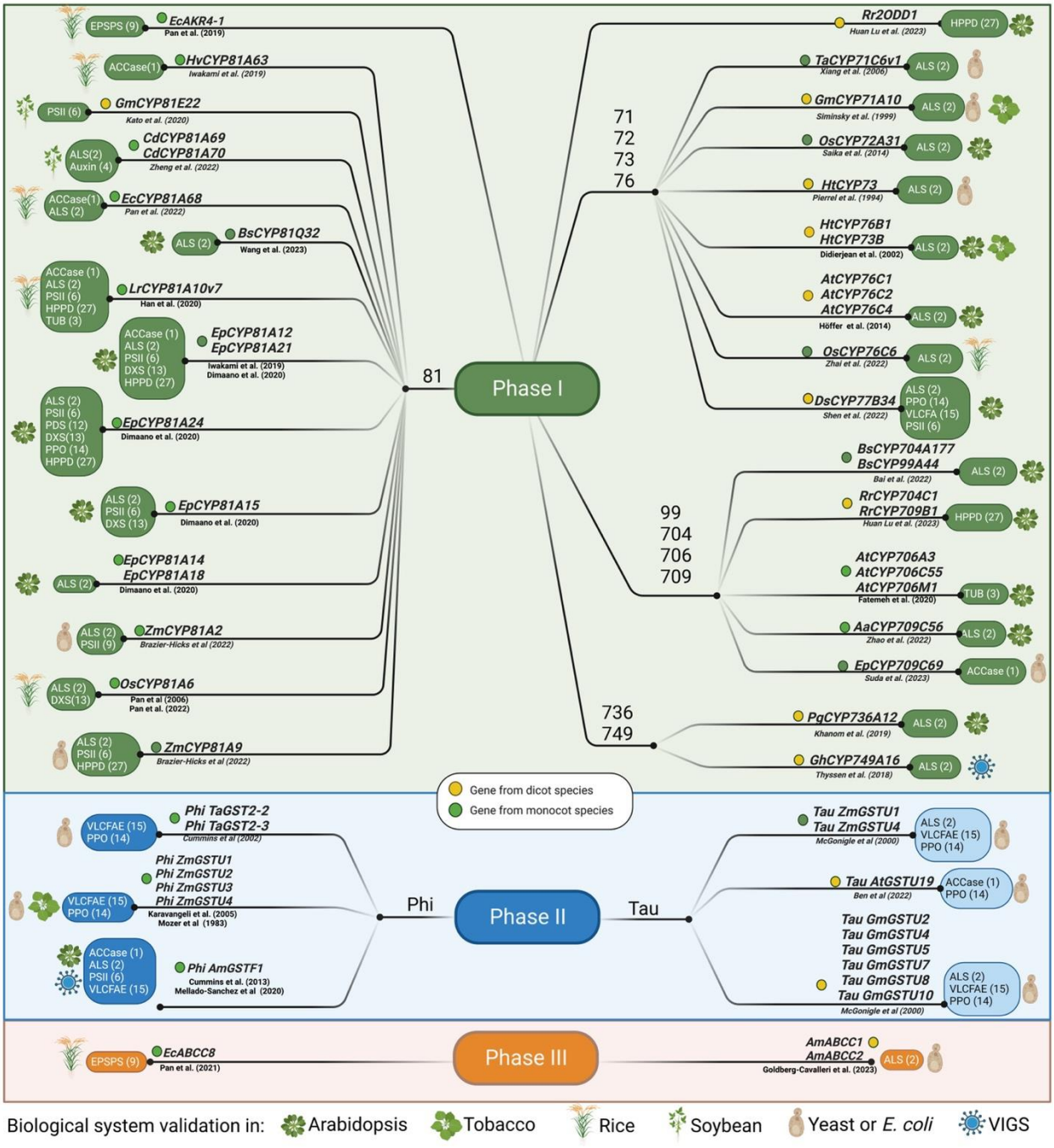


Figure 2-3. Illustration of herbicide metabolic resistance genes in weeds and their corresponding herbicide mechanisms of action (HRAC) for metabolism. Genes are separated by their function in the metabolic process, as phase I (activation), phase II (conjugation), and phase III (compartmentalization and incorporation). Green and yellow circles represent genes found in grass and broadleaf species, respectively. Note: VIGS - virus-induced gene silencing, Aa - *Alopecurus aequalis,* Am - *Alopecurus myosuroides*; At - *Arabidopsis thaliana*, Bs - *Beckmannia syzigachne*, Cd - *Cynodon dactylon*, Ds - *Descurainia sophia*, Ep - *Echinochloa phyllopogon*; Ec - *Echinochloa colona*; Gh - *Gossypium hirsutum*, Gm - *Glycine max*, Ht - *Helianthus tuberosus*; Hv - *Hordeum vulgare*, Lr - *Lolium rigidum*; Os - *Oryza sativa*, Pg - *Panax ginseng*, Rr -

Raphanus raphanistrum, Ta - *Triticum aestivum,* Zm - *Zea mays*. Genes: AKR=aldo-keto reductase, *CYP* - cytochrome P450 monooxygenase, *GST* - glutathione-S-transferase, *ABC* - ATP binding cassette.

REFERENCES

1. Peterson MA, Collavo A, Ovejero R, Shivrain V and Walsh MJ, The challenge of herbicide resistance around the world: a current summary. *Pest Manag Sci* **74**: 2246-2259 (2018).
2. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A and Dayan FE, Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry* **295**: 10307–10330 (2020).
3. Rigon CAG, Gaines TA, Küpper A and Dayan FE, Metabolism-based herbicide resistance, the major threat among the non-target site resistance mechanisms. *Outlooks on Pest Manag* **31**: 162-168 (2020).
4. Yu Q and Powles S, Metabolism-based herbicide resistance and cross-resistance in crop weeds: A threat to herbicide sustainability and global crop production. *Plant Physiol* **166**: 1106-1118 (2014).
5. Yuan JS, Tranel PJ and Stewart CN, Non-target-site herbicide resistance: a family business. *Trends Plant Sci* **12**: 6-13 (2007).
6. Edwards R and Dixon DP, Plant glutathione transferases. *Methods Enzymol* **401**: 169-186 (2005).
7. Switzer CM, The existence of 1,4-D resistant strains of wild carrot. *Proc NEWGC* **11**: 315-318 (1957).
8. Sweetser PB, Schow GS and Hutchison JM, Metabolism of chlorsulfuron by plants: Biological basis for selectivity of a new herbicide for cereals. *Pest Biochem Physiol* **17**: 18-23 (1982).
9. Georghiou G. Current problems, trends, and developments in pesticide metabolism in plants. In *Pesticide Resistance: strategies and tactics for management*. National Academy Press: Washington, DC, pp. 471 (1986).
10. Heap I and Knight R, A population of ryegrass tolerant to the herbicide diclofop-methyl [in South Australia]. *Journal of the Australian Institute of Agricultural Science* **48**1982).
11. Heap I and Knight R, The occurrence of herbicide cross-resistance in a population of annual ryegrass, *Lolium rigidum*, resistant to diclofop-methyl. *Australian Journal of Agricultural Research* **37**: 149-156 (1986).
12. kemp MS, Moss SR and Thomas TH. Herbicide resistance in *Alopecurus myosuroides*. In *Managing Resistance to Agrochemicals: from fundamentals research to practival strategies*, ed. by Green MB, LeBaron HM and Moberg WK. American Chemical Society; 1st edition, pp. 498 (1990).
13. LeBaron HM and Gressel J. *Herbicide resistance in plants*. John Wiley & Sons, Inc, New York, (1982).
14. Powles SB, Holtum JAM, Matthews JM and Liljegren DR. Herbicide Cross-Resistance in Annual Ryegrass *Lolium rigidum Gaud*. In *Managing Resistance to Agrochemicals*. American Chemical Society, pp. 394-406 (1990).
15. Burnet MWM, Loveys BR, Holtum JAM and Powles SB, Increased detoxification is a mechanism of simazine resistance in *Lolium rigidum*. *Pest Biochem Physiol* **46**: 207-218 (1993).
16. Christopher JT, Powles SB, Liljegren DR and Holtum JA, Cross-Resistance to Herbicides in Annual Ryegrass (Lolium rigidum) : II. Chlorsulfuron Resistance Involves a Wheat-Like Detoxification System. *Plant Physiol* **95**: 1036-1043 (1991).
17. Hall LM, Moss SR and Powles SB, Mechanism of Resistance to Chlorotoluron in Two Biotypes of the Grass Weed Alopecurus myosuroides. *Pest Biochem Physiol* **53**: 180-192 (1995).
18. Gronwald JW, Andersen RN and Yee C, Atrazine resistance in velvetleaf (Abutilon theophrasti) due to enhanced atrazine detoxification. *Pest Biochem Physiol* **34**: 149-163 (1989).
19. Manley BS, Hatzios KK and Wilson HP, Absorption, translocation, and metabolism of chlorimuron and nicosulfuron in imidazolinone-resistant and -susceptible smooth pigweed (*Amaranthus hybridus*). *Weed Technol* **13**: 759-764 (1999).

20. Cocker KM, Northcroft DS, Coleman JO and Moss SR, Resistance to ACCase-inhibiting herbicides and isoproturon in UK populations of Lolium multiflorum: mechanisms of resistance and implications for control. *Pest Manag Sci* **57**: 587-597 (2001).
21. Cocker KM, Coleman JOD, Blair AM, Clarke JH and Moss SR, Biochemical mechanisms of cross-resistance to aryloxyphenoxypropionate and cyclohexanedione herbicides in populations of Avena spp. *Weed Res* **40**: 323-334 (2000).
22. Park KW, Fandrich L and Mallory-Smith CA, Absorption, translocation, and metabolism of propoxycarbazone-sodium in ALS-inhibitor resistant Bromus tectorum biotypes. *Pest Biochem Physiol* **79**: 18-24 (2004).
23. Georghiou G and Taylor C. Factors influencing the evolution of resistance. In *Pesticide resistance*, ed. by Metcalf RL and WH L. National Academy Press: Washington D.C., pp. 157–169 (1986).
24. Neve P and Powles S, Recurrent selection with reduced herbicide rates results in the rapid evolution of herbicide resistance in *Lolium rigidum*. *Theoretical and Applied Genetics* **110**: 1154-1166 (2005).
25. Siminsky B, Plant cytochrome P450-mediated herbicide metabolism. *Phytochemistry Reviews* **49**: 445-448 (2006).
26. Yun MS, Yogo Y, Miura R, Yamasue Y and Fischer AJ, Cytochrome P-450 monooxygenase activity in herbicide-resistant and -susceptible late watergrass (*Echinochloa phyllopogon*). *Pest Biochem Physiol* **83**: 107-114 (2005).
27. Iwakami S, Endo M, Saika H, Okuno J, Nakamura N, Yokoyama M, Watanabe H, Toki S, Uchino A and Inamura T, Cytochrome P450 *CYP81A12* and *CYP81A21* are associated with resistance to two acetolactate synthase inhibitors in *Echinochloa phyllopogon*. *Plant Physiol* **165**: 618-629 (2014).
28. Dimaano NG, Yamaguchi T, Fukunishi K, Tominaga T and Iwakami S, Functional characterization of cytochrome P450 *CYP81A* subfamily to disclose the pattern of cross-resistance in *Echinochloa phyllopogon*. *Plant MolBiol* **102**: 403-416 (2020).
29. Gaines TA, Lorentz L, Figge A, Herrmann J, Maiwald F, Ott M-C, Han H, Busi R, Yu Q, Powles SB and Beffa R, RNA-Seq transcriptome analysis to identify genes involved in metabolism-based diclofop resistance in *Lolium rigidum*. *The Plant Journal* **78**: 865-876 (2014).
30. Duhoux A, Carrère S, Gouzy J, Bonin L and Délye C, RNA-Seq analysis of rye-grass transcriptomic response to an herbicide inhibiting acetolactate-synthase identifies transcripts linked to non-target-site-based resistance. *Plant Mol Biol* **87**: 473-487 (2015).
31. Zhao N, Yang J, Jiang M, Liao M and Cao H, Identification of essential genes involved in metabolism-based resistance mechanism to fenoxaprop-P-ethyl in Polypogon fugax. *Pest Manag Sci* **78**: 1164-1175 (2022).
32. Zhao N, Li W, Bai S, Guo WL, Yuan GH, Wang F, Liu WT and Wang JX, Transcriptome profiling to identify genes involved in mesosulfuron-methyl resistance in *Alopecurus aequalis*. *Front Plant Sci* **8**: 16 (2017).
33. Fang J, Zhang Y, Liu T, Yan B, Li J and Dong L, Target-Site and metabolic resistance mechanisms to penoxsulam in barnyardgrass (*Echinochloa crus-galli* (L.) P. Beauv). *J Agric Food Chem* **67**: 8085-8095 (2019).
34. Pan L, Yu Q, Han HP, Mao LF, Nyporko A, Fan LJ, Bai LY and Powles S, Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in *Echinochloa colona*. *Plant Physiol* **181**: 1519-1534 (2019).
35. Pan L, Yu Q, Wang J, Han H, Mao L, Nyporko A, Maguza A, Fan L, Bai L and Powles S, An ABCC-type transporter endowing glyphosate resistance in plants. *Proceedings of the National Academy of Sciences* **118**: e2100136118 (2021).
36. Murphy BP, Beffa R and Tranel PJ, Genetic architecture underlying HPPD-inhibitor resistance in a Nebraska *Amaranthus tuberculatus* population. *Pest Manag Sci* **77**: 4884-4891 (2021).

37. Kaundun SS, Hutchings SJ, Dale RP, Howell A, Morris JA, Kramer VC, Shivrain VK and McIndoe E, Mechanism of resistance to mesotrione in an *Amaranthus tuberculatus* population from Nebraska, USA. *PLoS One* **12**: 22 (2017).
38. Küpper A, Peter F, Zollner P, Lorentz L, Tranel PJ, Beffa R and Gaines TA, Tembotrione detoxification in 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor-resistant Palmer amaranth (*Amaranthus palmeri* S. Wats.). *Pest Manag Sci* **74**: 2325-2334 (2018).
39. Lu H, Yu Q, Han H, Owen MJ and Powles SB, Evolution of resistance to HPPD-inhibiting herbicides in a wild radish population via enhanced herbicide metabolism. *Pest Manag Sci* **n/a**2019).
40. Busi R, Girotto M and Powles SB, Response to low-dose herbicide selection in self-pollinated *Avena fatua*. *Pest Manag Sci* **72**: 603-608 (2016).
41. Suda H, Kubo T, Yoshimoto Y, Tanaka K, Tanaka S, Uchino A, Azuma S, Hattori M, Yamaguchi T, Miyashita M, Tominaga T and Iwakami S, Transcriptionally linked simultaneous overexpression of P450 genes for broad-spectrum herbicide resistance. *Plant Physiol* 2023).
42. Werck-Reichhart D, Bak S and Paquette S, Cytochromes p450. *Arabidopsis Book* **1**: e0028-e0028 (2002).
43. Wei K and Chen H, Global identification, structural analysis and expression characterization of cytochrome P450 monooxygenase superfamily in rice. *BMC Genomics* **19**: 35 (2018).
44. Wu D, Shen E, Jiang B, Feng Y, Tang W, Lao S, Jia L, Lin H-Y, Xie L, Weng X, Dong C, Qian Q, Lin F, Xu H, Lu H, Cutti L, Chen H, Deng S, Guo L, Chuah T-S, Song B-K, Scarabel L, Qiu J, Zhu Q-H, Yu Q, Timko MP, Yamaguchi H, Merotto A, Qiu Y, Olsen KM, Fan L and Ye C-Y, Genomic insights into the evolution of Echinochloa species as weed and orphan crop. *Nat Commun* **13**: 689 (2022).
45. Saika H, Horita J, Taguchi-Shiobara F, Nonaka S, Nishizawa-Yokoi A, Iwakami S, Hori K, Matsumoto T, Tanaka T, Itoh T, Yano M, Kaku K, Shimizu T and Toki S, A novel Rrce cytochrome P450 gene, CYP72A31, confers tolerance to acetolactate synthase-inhibiting herbicides in rice and arabidopsis. *Plant Physiol* **166**: 1232-+ (2014).
46. Pan G, Zhang XY, Liu KD, Zhang JW, Wu XZ, Zhu J and Tu JM, Map-based cloning of a novel rice cytochrome P450 gene *CYP81A6* that confers resistance to two different classes of herbicides. *Plant MolBiol* **61**: 933-943 (2006).
47. Guo F, Endo M, Yamaguchi T, Uchino A, Sunohara Y, Matsumoto H and Iwakami S, Investigation of clomazone-tolerance mechanism in a long-grain cultivar of rice. *Pest Manag Sci* **77**: 2454-2461 (2021).
48. Kato S, Yokota Y, Suzuki R, Fujisawa Y, Sayama T, Kaga A, Anai T, Komatsu K, Oki N, Kikuchi A and Ishimoto M, Identification of a cytochrome P450 hydroxylase, *CYP81E22*, as a causative gene for the high sensitivity of soybean to herbicide bentazon. *Theoretical and Applied Genetics* **133**: 2105-2115 (2020).
49. Siminszky B, Corbin FT, Ward ER, Fleischmann TJ and Dewey RE, Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast and tobacco enhances the metabolism of phenylurea herbicides. *Proc Natl Acad Sci U S A* **96**: 1750-1755 (1999).
50. Brazier-Hicks M, Franco-Ortega S, Watson P, Rougemont B, Cohn J, Dale R, Hawkes TR, Goldberg-Cavalleri A, Onkokesung N and Edwards R, Characterization of cytochrome P450s with key roles in determining herbicide selectivity in maize. *ACS Omega* 2022).
51. Zheng T, Yu X, Sun Y, Zhang Q, Zhang X, Tang M, Lin C and Shen Z, Expression of a cytochrome P450 gene from Bermuda grass *Cynodon dactylon* in soybean confers tolerance to multiple herbicides. *Plants* **11**: 949 (2022).
52. Han H, Yu Q, Beffa R, González S, Maiwald F, Wang J and Powles SB, Cytochrome P450 *CYP81A10v7* in *Lolium rigidum* confers metabolic resistance to herbicides across at least five modes of action. *The Plant Journal* **105**: 79-92 (2021).

53. Wang J, Lian L, Qi J, Fang Y, Nyporko A, Yu Q, Bai L and Pan L, Metabolic resistance to acetolactate synthase inhibitors in *Beckmannia syzigachne*: identification of *CYP81Q32* and its transcription regulation. *The Plant Journal* **n/a**2023).
54. Bai S, Yin M, Lyu Q, Jiang B and Li L, Cytochrome P450 *BsCYP99A44* and *BsCYP704A177* confer metabolic resistance to ALS herbicides in *Beckmannia syzigachne*. *Int J Mol Sci* **23**: 12175 (2022).
55. Shen J, Yang Q, Hao L, Zhang L, Li X and Zheng M, The metabolism of a novel cytochrome P450 (*CYP77B34*) in tribenuron-methyl-resistant *Descurainia sophia* L. to herbicides with different mode of actions. *Int J Mol Sci* **23**: 5812 (2022).
56. Lu H, Liu Y, Li M, Han H, Zhou F, Nyporko A, Yu Q, Qiang S and Powles S, Multiple metabolic enzymes can be involved in cross-resistance to 4-hydroxyphenylpyruvate-dioxygenase-inhibiting herbicides in wild radish. *J Agric Food Chem* 2023).
57. Mozer TJ, Tiemeier DC and Jaworski EG, Purification and characterization of corn glutathione S-transferase. *Biochemistry* **22**: 1068-1072 (1983).
58. McGonigle B, Keeler SJ, Lau SM, Koeppe MK and O'Keefe DP, A genomics approach to the comprehensive analysis of the glutathione S-transferase gene family in soybean and maize. *Plant Physiol* **124**: 1105-1120 (2000).
59. Karavangeli M, Labrou NE, Clonis YD and Tsaftaris A, Development of transgenic tobacco plants overexpressing maize glutathione *S*-transferase I for chloroacetanilide herbicides phytoremediation. *Biomol Eng* **22**: 121-128 (2005).
60. Cummins I, Wortley DJ, Sabbadin F, He ZS, Coxon CR, Straker HE, Sellars JD, Knight K, Edwards L, Hughes D, Kaundun SS, Hutchings SJ, Steel PG and Edwards R, Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. *Proceedings of the National Academy of Sciences* **110**: 5812-5817 (2013).
61. Mellado-Sánchez M, McDiarmid F, Cardoso V, Kanyuka K and MacGregor DR, Virus-mediated transient expression techniques enable genetic modification of *Alopecurus myosuroides*. *bioRxiv*: 2020.2001.2028.923466 (2020).
62. Goldberg-Cavalleri A, Onkokesung N, Franco-Ortega S and Edwards R, ABC transporters linked to multiple herbicide resistance in blackgrass (*Alopecurus myosuroides*). *Front Plant Sci* **14**: 1082761 (2023).
63. González-Torralva F, Brown AP and Chivasa S, Comparative proteomic analysis of horseweed (*Conyza canadensis*) biotypes identifies candidate proteins for glyphosate resistance. *Sci Rep* **7**: 42565 (2017).
64. Hikosaka M, Iwahashi F and Yamato S, Metabolomic analysis of Schoenoplectus juncoides reveals common markers of acetolactate synthase inhibition among paddy weeds. *Pest Biochem Physiol* **174**: 104827 (2021).
65. Dong H, Huang Y and Wang K, The Development of Herbicide Resistance Crop Plants Using CRISPR/Cas9-Mediated Gene Editing. *Genes (Basel)* **12**2021).
66. Zabala-Pardo D, Gaines T, Lamego FP and Avila LA, RNAi as a tool for weed management: challenges and opportunities. **40**: - (2022).

3. CHAPTER II - CYTOCHROME *CYP72A219* IS INVOLVED IN METABOLIC RESISTANCE TO TEMBOTRIONE IN PALMER AMARANTH

INTRODUCTION

The emergence and spread of herbicide-resistant weeds poses a major challenge to sustainable agriculture, as they can diminish crop yields and increase the costs associated with weed control[1, 2]. Evolution of herbicide resistance is a consequence of consistent selection pressure exerted by the repeated use of herbicides with the same mode of action over time[3]. This selection process gives rise to herbicide-resistant plants through different mechanisms, primarily classified as target-site (TSR) and non-target-site resistance (NTSR)[3, 4]. Between these two groups, NTSR, specifically metabolic resistance, is an especial critical issue due to the potential interaction of a single metabolic resistance mechanism with a wide range of different classes of herbicides[4].

Plants possess a diverse array of enzymes involved in the detoxification of xenobiotics. Within these enzymes, cytochrome P450 enzymes (P450) play a key role in the initial step of the detoxification process, while GST and GT enzymes function in the subsequent steps, and ABC transporters facilitate compartmentalization[3, 4]. These enzymes have been identified in various crops, demonstrating their ability to metabolize herbicides in heterologous systems, including CYP81E22 and CYP71A10 from soybean[5, 6], CYP81A6 from rice[7, 8], CYP81A9 from corn[9], GST2 from wheat[10] and GSTU from corn[11]. Over the past decade, there has been an increasing identification of genes in resistant weeds capable of metabolizing different herbicides. Notable examples include *CYP81A12*, *CYP81A21*, *CYP81A15*, and *CYP81A14* from *Echinochloa phyllopogon*[12-14], *CYP81A69* and *CYP81A70* from *Cynodon dactylon*[15], *CYP81A10v7* from *Lolium rigidum*[16], *CYP77B34* from *Descurainia sophia*[17], *CYP736A12* from *Panax ginseng*[18],

GSTF1, *ABCC1* and *ABCC2* from *Alopecurus myosuroides*, and *ABCC8* from *Echinochloa colona*. Additionally, other types of oxidoreductase enzymes have been identified to play a role in herbicide resistance, such as AKR4-1 from *Echinochloa colona* which metabolizes glyphosate into less toxic metabolites[19], *HIS1* from rice that encodes a 2-oxoglutarate (2OG)/Fe(II)-type oxidoreductase capable of detoxifying triketone by hydroxylation and recently ODD1, encoding 2-oxoglutarate/Fe(II)-dependent dioxygenase, *from Raphanus raphanistrum*, providing resistance to tembotrione and isoxaflutole[20]. The majority of the identified genes are found in grasses, with only a few genes identified in dicot weeds[4].

Palmer amaranth (*Amaranthus palmeri*) is a highly invasive and troublesome weed that poses a significant threat to agricultural crops and ecosystems. Native to the southwestern United States, this annual broadleaf weed has become a pervasive problem across various regions due to its aggressive growth, adaptability, and resistance to herbicides[21, 22]. Its dioecious nature, with separate male and female plants, allows for efficient cross-pollination, leading to the exchange of genetic material and the emergence of novel traits[21]. Combined with its fast growth and extensive root system, Palmer amaranth possesses a remarkable capacity for adaptation, including the evolution of resistance to many different herbicides. Palmer amaranth populations evolved resistance to eight different modes of action, including those targeting acetolactate synthase (ALS)[23, 24], 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)[25, 26], synthetic auxins [27], protoporphyrinogen oxidase (PPO)[28], photosystem II (PSII)[29], 4-hydroxyphenylpyruvate dioxygenase (HPPD)[29, 30], very long chain fatty acid elongase (VLCFAE)[31], and tubulin [32].

Palmer amaranth has evolved herbicide resistance through two main mechanisms, TSR and NTSR. The most commonly observed cases of resistance involve TSR to ALS- and EPSPS-

inhibitors. However, there is an increasing resistance cases due to metabolic mechanism[27, 30], and further analysis is necessary to identify the specific genes involved and their regulation. RNA-seq analysis plays a crucial role in deciphering the genetic basis of herbicide resistance, especially if it is a multigenic trait[33, 34]. Additionally, with the availability of complete genomes for weed species, such as Palmer amaranth, conducting QTL (Quantitative Trait Locus) mapping experiments becomes a valuable approach for mapping key traits in resistant weeds[35]. By combining the comprehensive gene expression data obtained from RNA-seq analysis with QTL mapping, researchers can gain valuable insights into the genetic factors underlying resistance and other important characteristics in weeds.

The research aimed to achieve several objectives. Firstly, it sought to identify and validate the metabolic genes associated with the tembotrione-resistance mechanism in a specific population of Palmer amaranth that exhibited enhanced metabolism. Additionally, the study aimed to determine if these identified causal genes could also be involved in conferring resistance in other Palmer amaranth populations that exhibit enhanced metabolism towards HPPD inhibitors. Finally, the research aimed to deepen our understanding of the regulatory mechanisms governing the identified genes responsible for herbicide resistance.

MATERIALS AND METHODS

Plant material for QTL mapping and RNA-seq analysis

The resistant (NER) and susceptible (NES) *A. palmeri* populations were collected from fields in Shickley, Nebraska in 2011. NER is resistant to atrazine and the HPPD inhibitors tembotrione, mesotrione, and topramezone [29, 30]. The enhanced metabolism of tembotrione to hydroxy-tembotrione was identified as the mechanism of herbicide resistant in NER population [30]. *A. palmeri* is a species with high genetic variability. Therefore, plants with a similar genetic

background were generated for QTL mapping and RNA-seq experiment by controlled pairings of NER and NES parents to utilize genetic recombination in an F2 population to minimize the impact of genetic differences unrelated to EMR traits. The approaches for the generation of the Pseudo-F2 plants are in Appendix A Supporting Information Material and Methods and Supporting Information Figure 3-1.

RNA-sequencing analysis

Descriptions of plant population, herbicide application, plant tissue collection, and library preparation are in Appendix A Supporting Information Material and Methods. Raw reads were pre-processed by removal of library adapter sequences, removal of low quality reads, and assessing the quality control using fastp [36]. Read alignment of the 108 libraries to the available male genome of Palmer amaranth (v1.1, id55760) [37] was performed using Hisat2.2.1 [38]. Most of the reads (>77%) aligned concordantly once, while >7% aligned concordantly more than one time and about 14% of reads aligned not concordantly. Mapped reads were assigned using featureCounts [39].

P450 gene sequences and phylogenetic tree

Consensus sequences of the candidate P450 genes were extracted from the transcriptome and aligned to the reference Palmer amaranth genome[37] to analyze SNPs. PCR was performed to amplify the coding sequences in four samples of each S and R from pseudo-F2 population that were used for the RNA-seq experiment. cDNA synthesis was performed using ProtoScript® II First Strand cDNA synthesis kit with 1 ug of RNA and treatment with DNase I. Protein sequences from cytochrome P450 known to metabolize herbicide in different weed species were obtained from NCBI. Multiple protein alignment was performed using ClustalO and used for tree

construction with Neighbor-joining method using Geneious Prime® 2023.0.1. P450 sequences are listed in Supporting Information.

Quantitative reverse transcription-PCR (RT-qPCR) validation

NES and NER populations were grown and herbicide applied at 91 g a.i. ha^{-1} as mentioned before. Four plants of each population were used for gene expression analysis validation. Youngest leaf tissue was collected at 0, 3, 6, and 12 h after herbicide application in 2 mL Eppendorf tubes and placed in liquid nitrogen. The tubes were kept at -80 ºC for further analysis. Tissue was ground with 3 mm stainless steel beads in a TissueLyser (Qiagen) with intensity of 30 for 1 min. RNA was isolated using Direct-zol RNA Miniprep from Zymo Research. cDNA synthesis was performed using ProtoScript® II First Strand cDNA synthesis kit using 1 µg of RNA and purified with DNase I.

Relative gene expression was analyzed on a T100 Thermal Cycler (BioRad), using SsoAdvanced™ universal SYBR® Green supermix. Reactions mixtures consisted of 10 µL of SsoAdvanced universal SYBR® Green supermix (2×), 2.5 µL of forward and reverse primers at 10 µM and 5 µL of cDNA (1:20). Thermocycler conditions consisted of an initial step of 30 sec at 95°C followed by 35 cycles of 5s at 94C and 30 s at 60C. Melting curve analysis was added using 65°C with 0.5°C increments of 5 sec/step.

The reference genes used were *18S rRNA* and *Actin7*. These two genes had the best gene stability as assessed by the NormFinder algorithm [40] (Supporting Information Table 3-1). Primers for the reference genes were designed based on conserved regions after alignment of *18S (FJ669720.1), actin* (HQ656028.1) and *TUB* (XM_010693569.3) from *Beta vulgaris* with the reference genome from Palmer amaranth (CoGe – v1.1, id55760). The candidate genes tested based on RNA-seq results were *CYP72A219*_4284, *CYP72A219*_4285, *CYP72A219*_7285, and

CYP81E8. Another *CYP72A-like* gene, hereafter named as *CYP72A219*_4286 was used as a cytochrome P450 not associated with herbicide resistance based on RNA-seq data. This gene is localized around 1200 bp downstream of the gene *CYP72A219*_4285 and in antisense strand based on reference genome. Primers for candidate genes were designed on regions based on consensus sequences from the transcriptome and Sanger sequencing of the genes. Primers sequences are listed in Supporting Information Table 3-2.

Heterologous expression of CYPs

To explore the potential involvement of candidate genes, which exhibited constitutive up-regulation in the RNA-seq data, in tembotrione metabolism, a heterologous expression approach was employed. The genes of interest were introduced and expressed in yeast for further investigation of their role in tembotrione metabolism. Coding sequences of the candidate genes were sent for gene optimization for yeast transformation and inserted in the pUC-WG/ampv vector. The gene sequences used were *CYP72A219*_4284 allele 1 and 2, *CYP72A219_4285*, *CYP72A219_7285* allele 1 and 2, and *CYP81E*8 allele 1 and 2. Another version of *CYP81E8*, named as *CYP81E8_v.18aa*, was synthesized with an insertion of 54 bp in the position 191 of the gene which corresponds to a sequence that is present in cytochrome P450 genes of the family *CYP81* from grasses (Supporting Information). As positive control, the gene *CYP81A9* from corn, known as *Nsf1*, was used. This gene was identified as the major resistance locus in corn for the herbicides nicosulfuron and mesotrione [41]. Restriction sites for BamHI and EcoRI were added the the 5' and 3' ends, respectively. Kozak sequence (AAAAAATCT) was added at the 5' end as a protein translation initiation site. Restriction enzyme reactions were performed to cut the optimized sequence from the vector using 1 µL EcoRI, 1 µL BamHI, 500 ng of the vector, with

incubation at 37°C for 2 h. Reaction products were run 1% agarose gel electrophoresis and the gene band was cut and purified using DNA Gel Purification Kit from New England Biolabs®.

The pYES2 yeast expression vector was used for recombinant expression. It contains the *URA3* gene for selection in yeast and 2μ origin for high-copy maintenance and GAL1 promoter to express protein (Thermo Fisher Scientific). The ligation reaction was performed using LigaFast™ Rapid DNA Ligation System from Promega. The reaction consisted of 50 ng digested pYES2 plasmid vector, 5 μL 2× Rapid Ligation Buffer, 1 μL T4 DNA ligase (3 Weiss unit/μL), 50 ng of gene and purified water up to 10 μL. The reaction was incubated overnight at 4°C. The product from the reaction was used to transform *E. coli* using the One Shot® TOP 10 kit (Invitrogen). The transformed cells were plated in petri dish with ampicillin and incubated at 37°C overnight. Single colonies were selected, and gene insertion was confirmed by colony PCR using high-fidelity PrimerStar HS DNA polymerase. The gene sequence was confirmed by Sanger sequencing. Plasmids were isolated using ZymoPure II Plasmid Miniprep Kit.

Yeast transformation and tembotrione incubation

The strains WAT11 and WAT21 of *Saccharomyces cerevisiae* containing cytochrome P450 reductase 1 and 2 [42], respectively, were used as a heterologous system to test the hypothesis that the candidate P450 genes could metabolize tembotrione. Yeast cells were grown in glucose SC (-ura) agar plates. Single cells were inoculated in YPD(A) medium and transformed using a modified lithium acetate procedure [43]. Transformed yeast cells were selected by glucose SC (-ura) agar plates. Yeast colony PCR to confirm the gene insertion was performed using high-fidelity PrimerStar HS DNA polymerase by heating the master mix at 94°C for 4 min before the PCR protocol.

Single colonies containing empty pYES2 or the candidate CYP genes were incubated in 15 mL 2% raffinose medium and allowed to grow for 2-3 d at 30 °C. Cell density was measured spectrophotometrically until the OD600 in 20 mL of induction medium was 5. The exact volume was removed and pelleted at 1,500 × g for 15 min at 4 °C. The cells were resuspended with 1 mL of induction medium containing 2% galactose and inoculated into 20 mL of the same medium. Tembotrione at 1500 µM diluted in EtOH was applied right after the cells were incubated in the induction medium. The yeast cells were incubated at 30°C at 200 rpm. Twenty-four hours after herbicide application, 5 mL of the medium were collected in 15 mL falcon tubes and cells were pelleted at 1,500 × g for 5 min. The supernatant was collected, and cleaned by adding 5 mL of 5% acetic acid acetonitrile + quenchers, vortexed for 30 s, and centrifuged for 15 min at 2,000 × g. The supernatant was collected, filtered through an Econofltr Nyln 13 mm 0.2µm (Agilent Technologies), and injected in the LC/MSMS.

LC-MSMS protocol

LC-MS/MS system consisted of a Nexera X2 UPLC with 2 LC-30AD pumps, a SIL-30AC MP autosampler, a DGU-20A5 Prominence degasser, a CTO-30A column oven, and SPD-M30A diode array detector coupled to an 8040 quadrupole mass-spectrometer. For tembotrione, the MS was in negative mode with a MRM optimized for 439.1>226.05 and set for 100 ms dwell time with a Q1 pre-bias of 11.0V, a collision energy of 11.0V and a Q3 pre-bias of 14.0V. For hydroxy-tembotrione, the MS was in negative mode with a MRM optimized for 455.1>419.05 and set for 100 ms dwell time with a Q1 pre-bias of 11.0V, a collision energy of 11.0V and a Q3 pre-bias of 14.0V. The samples were chromatographed on a 100×4.6 mm Phenomenex Kinetex 2.6 µm biphenyl column maintained at 40°C. Solvent A consisted of water with 0.1% formic acid and solvent B was acetonitrile with 0.1% formic acid. The solvent program started at 80% B

and increased to 100% B in 3.5 min and maintained at 100% for 2 min. The solvent was returned to 80% B and maintained there for 3 min before the next injection. The flow rate was set at 0.4 mL/min and samples were analyzed as 1 µL injection volumes.

Promoter sequencing and validation

CYP72A219_4284 and CYP81E8 promoter amplification

To investigate the potential regulatory mechanisms governing the expression of candidate genes, the promoter region was sequenced. The aim was to identify any *cis* elements present within the promoter that could potentially influence gene expression. Pseudo-F2 plants were grown in the greenhouse and the youngest leaf tissue was collected when the plants achieved four to five true leaf stage. Tembotrione was applied at 77 g ha^{-1} rate. Shoot fresh weight was measured at 28 d after herbicide application. Five resistant and susceptible plant from a pseudo-F2 population produced by crossing NER and NES were chosen for promoter amplification. DNA extraction was performed using modified hexadecyltrimethylammonium bromide (CTAB) method. DNA quantification was carried out using nanodrop 2000c. Primers were designed by Primer3Plus (https://www.bioinformatics.nl/) using as reference the available draft genome of Palmer amaranth [37]. The primers sequences used to amplify the genes are listed in the Supporting Information Table 3-2. Reverse primers were designed on the conserved sequence of the first exon for each gene. PCR was performed using PrimeSTAR® HS DNA Polymerase kit (Takara) consisting of 10 µL 5× PrimSTAR buffer (Mg2+ Plus), 4 µL dNTP mix (2.5 mM each), 1.5 µL of each primer (forward and reverse), 0.5 µL PrimeSTAR HS DNA Polymerase, 1 µL of DNA (50 ng) and sterile water up to 50 µL. PCR cycling conditions were initial denaturation at 98C for 30 s, followed by 40 cycles of denaturation at 98°C for 10 s, annealing at 60°C for 15 s and extension at 72°C for 2.5 min. PCR product was run in 1% agarose gel electrophoresis for 30

min to check for a single band. Amplicons were sent for sequencing using the long-read sequencing technology Oxford Nanopore Technology (ONT) by SNPsaurus LLC (https://www.plasmidsaurus.com). Sequences results were confirmed by blasting the reads to the Palmer amaranth genome and submitted to analysis for common or different motifs using the Multiple Expectation maximizations for Motif Elicitation (MEME-suite) tool [44]. Nucleotide motifs were scanned for biological roles using Gene Ontology for Motifs (GOMo) tool [45] to determine if any motif was significantly associated with genes linked to one or more genome ontology (Go) using *Arabidopsis thaliana* database.

Construct of promoter::GUS vector

For the functional validation of the *CYP72A219* promoter from R and S Palmer amaranth pseudo-F2 population, three 5'- fragments composing different lengths of the *CYP72A219_4284* gene promoter (–1500 bp, – 750bp and –250bp) from sensitive and resistant plants were cloned into vector pBI121 replacing the *35S::GUS* promoter. Nested PCR reaction was performed to amplify the different promoter lengths using product PCR cited above (Figure 3-1A). The pairs of primers were designed specifically for each population. A set of primers was designed to amplify a chunk DNA from *Arabidopsis thaliana* as a negative control. Primers were designed by adding a leader sequence, BamHI and SbfI restriction site in the reverse and forward primers, respectively. The primers sequences and their composition are listed in the Supporting Information Table 3-2. The PCR product was run in 1% agarose gel electrophoresis to confirm the different promoter sizes amplification. Digestion reaction was performed for each of the treatments and vector pBI121 using 1 µg of DNA, 5 µL of 10× rCutSmart buffer, 1 µL of each restriction enzyme SbfI-HF and BamHI-HF, and nuclease-free water to 50 µL. The samples were incubated for 15 min at RT. 6× DNA loading dye was added into the samples and run in 1%

agarose gel electrophoresis (Figure 3-1B), bands were cut and purified using Monarch® DNA Gel Extraction kit.

Ligation reaction was performed using 1 µL T4 ligase PROMEGA, 5 µL 2× rapid ligation buffer, 50 ng vector and insert in a proportion of 3:1. Reactions were incubated in room temperature for 15 min. Transformation of *E. coli* was performed using One Shot™ TOP10 Chemically Competent cells following the manual instructions. Cells were plated in selection medium containing kanamycin. Colony PCR was performed for six colonies for each construct using primers pBI121_F, designed on *GUS* coding sequence and M13_R (Supporting Information Table 3-2) to check for correct replacement of 35S promoter by the ones from *CYP72A219_4284*. Transformants were incubated in LB broth medium, and plasmid were extracted using ZymoPURE plasmid minipred kit. Plasmid were sent for full-length sequencing and annotation of clonal circular plasmid DNA using long-read sequencing technology from Oxford Nanopore Technologies (ONT) by SNPsaurus LLC (https://www.plasmidsaurus.com). The different vectors achieved were *35S::GUS* (positive control), *chunk::GUS* (negative control), *250_S::GUS* (250 bp promoter from Pseudo-F2 sensitive plants), *250_R::GUS* (250 bp promoter from Pseudo-F2 resistant plants), *750_S::GUS*, *750_R::GUS*, *1500_S::GUS* and *1500_R::GUS*.

Agrobacterium tumefaciens transformation

The heat/shock method was used to transform *Agrobacterium* with the different constructs of promoter::GUS. *Agrobacterium* strain GV3101 was inoculated in 3 mL LB supplemented with gentamycin (50mg/L) and kanamycin (50mg/L) in a 15 mL cell culture tubes and grown overnight at 30 °C. Cells were inoculated into 50 mL LB medium and allowed to grow until it reaches OD600 of 0.8. Cells were pelleted by centrifugation at 2,000 × g for 5 min and kept at

4°C. The supernatant was discarded, and pellet was resuspended with 1 mL of ice cold 20 mM $CaCl_2$. In pre-chilled 1.5 mL tubes, 0.1 mL of bacterial suspension was dispensed and combined with 1 µg DNA of each vector with gentle mixing. The control treatment had no addition of DNA. The transformations were performed by freezing the tubes in liquid nitrogen and thawing for 5 min at 37°C, followed by adding 1 mL LB medium and incubation for 2 h at 28°C and shaking at 10 × g. The *Agrobacterium* cells were transformed with plasmid pBI121 that harbors a *GUS* gene under the control of different promoter, 35S CaMV (positive control), DNA chunk (negative control), 250_S, 250_R, 750_S, 750_R, 1500_S, 1500_R. The transformants were plated on LB plates with kanamycin and gentamycin as selection marker and incubated for 2 d at 28°C. Colony PCR was performed using the same primers indicated above to confirm the transformation.

Tobacco growth condition and Agroinfiltration

Seeds of *Nicotiana benthamiana* were planted into the pots containing soil in growth chamber under the conditions of 50%–75% relative humidity, 16 h light/8 h darkness photocycle at 23°C. The 4-week-old plants were used for agroinfiltration. Agroinfiltration was performed as described by Liu et al.[46]. *Agrobacterium* strain GV3031 containing each individual construct described earlier was incubated in 2 mL LB medium supplemented with appropriate antibiotics (kanamycin and gentamycin) and then inoculated (1%) into 25 mL LB with 10 mM MES, 20 µM acetosyringone as well as the antibiotics. The culture was grown to achieve log phase (OD600 0.8) at 28°C, centrifuged and resuspended in MMA solution (10 mM MES, 10 mM $MgCl_2$, 100 µM acetosyringone) to a final OD600 of 0.8 and kept at room temperature for 3 h. One mL of the bacterial suspension was infiltrated into intercellular spaces of fully expanded leaves using a 1 mL plastic syringe. Typically, up to four infiltrating spots separated by veins could be arranged

in a single leaf of *N. benthamiana*. Each treatment was infiltrated in three different leaves and in three different plants. Infiltrated area was marked for later analysis. After agroinfiltration, the treated plants were maintained in the growth chamber at the dark for 2 d and samples for GUS analysis were collected 4 d after the infiltration.

GUS expression and histochemical assay

For *GUS* expression analysis, in the infiltrated area, two leaf discs (1.3 cm diameter) were collected for each leaf, totaling a biological replicate with six discs, two of each treated leaf. Samples were collected in 1.5 mL tubes and kept in liquid nitrogen and subsequently stored at -80°C. RNA extraction and cDNA synthesis were performed as described above. Reference genes evaluated were *18S* rRNA (accession number TC23401 - AtMg01390) and *GAPDH* (accession number TC21175 - At1g12900) and its primers sequences were based on previous studies [47]. Primer set for *GUS* gene were designed based on gene sequence in pBI121 vector sequences results. Primers sequences are listed in the Supporting Information Table 3-2.

For the GUS histochemical analysis, one leaf disc (1.6 cm diameter) was collected from each treated leaf, totaling three leaf discs for each plant replicate and incubated in 6-well culture plates with 2 mL of 50 mM Na_3PO_4 buffer at pH 7.2, 10 mM EDTA, 2 mM $K_3Fe(CN)_6$, 2 mM $K_4Fe(CN)_6$ and 1mM of X-gluc. The samples were vacuum infiltrated for 1.5 min and incubated at 37°C overnight in the dark. The solution was removed, and leaf discs were incubated with 70% ethanol shaking at 150 rpm. Ethanol was replaced by 80% ethanol until all chlorophyll was removed. Discs were incubated with 70% ethanol until assayed.

Involvement of *CYP72A219_4284* in different HPPD-resistant Palmer amaranth populations

Seeds from Palmer amaranth populations were collected in the season 2019 in United States with suspected herbicide resistance to mesotrione and tembotrione. Initial screening was performed in the greenhouse using one and two times the label rate of mesotrione and tembotrione to confirm the herbicide resistance. Ten suspected resistant populations (WR2019-274, WR2019-141, WR2019-140, WR2019-137, WR2019-199, WR2019-200, WR2019-044, WR2019-144, WR2019-198, WR2019-273), two known tembotrione-resistant (WR2013-034 and NER) and one sensitive control (IHX_3361) population were used for the following experiments.

Whole-plant dose response

Seeds from 13 Palmer amaranth populations were sown in plastic trays filled with soil and kept in the greenhouse at 28º C and photoperiod of 16 h light. After one week, two seedlings were transplanted into single fiber pots with dimensions of 11 × 7 × 11.5 composing one replicate. The dose response experiment consisted of 11 doses, which were 0, 1/128, 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 1, 2 and 4 × the label rate (91 g a.i. ha^{-1}). Every dose had six replicates, with a total of 12 plants for each dose. The herbicide tembotrione (Laudis, 419 g a.i. L^{-1}, Bayer, Leverkusen, Germany) was applied together with 2,200 g a.i. ha^{-1} of the wetting agent Mero (Bayer, Leverkusen, Germany) and 170 g a.i. ha^{-1} ammonium sulfate using a stationary research sprayer (Höchst AG, Höchst, Germany) calibrated to deliver a spray volume of 300 L ha^{-1}. Survival and fresh shoot weight were recorded 28 d after application.

^{14}C-Tembotrione metabolism

Metabolism of tembotrione was measured in 13 populations over time in six individuals per population and treatment. Tembotrione was applied on the two youngest expanded leaves of individuals at the four-leaf stage with a total of ten 1-µL droplets (5 µL per leaf) of ^{14}C-tembotrione (Bayer, Leverkusen, Germany) in a 0.3% v/v Mero solution (Bayer) with 3.3 kBq or 200,000 dpm μL^{-1}, corresponding to 0.762 µg $\mu L^{-1}$of tembotrione. Treated plants were kept in a growth chamber at 28°C under continuous light conditions with a light intensity of 500 µmol m^{-2} s^{-1} and 70% humidity. Plant shoots were harvested at 6, 12, 24, and 48 h after treatment (HAT).

The harvested tissue was washed in 80% acetone three times to remove any non-absorbed ^{14}C-tembotrione, and then disrupted in 500 µL of methanol with 5-mm stainless steel beads at 30 Hz for 10 min. The homogenate was centrifuged at 6,000 × g for 10 min. The residue was re-extracted with 600 µL of methanol followed by a final extraction with 600 µL of 90% acetonitrile. All solvents used were high-performance liquid chromatography (HPLC) grade (Sigma-Aldrich, Steinheim, Germany; ≥ 99.9 % HPLC grade).

The pooled supernatant was evaporated under continuous air flow at 55 °C and then re-suspended in 200 µL of 90% acetonitrile using a shaker and ultrasonic bath and then filtered through a 0.45-µm low-binding hydrophilic polytetrafluoroethylene (PTFE) mesh for 10 min at 2,200 × g in the centrifuge. The recovered radioactivity in the filtrate was 92% of the total applied, on average.

A non-treated control sample, spiked with ^{14}C-tembotrione just prior to extraction, was also included. Separation and HPLC identification of the parent tembotrione herbicide and its metabolites were performed on a reverse-phase HPLC system (LC Net II/ADC with PU-980 pump unit, LC-980-02 gradient unit and CO-2060 Plus column thermostat; Jasco, Oklahoma

City, OK, USA). Chromatographic separation was achieved with a 150 × 2.0 or 3.0 mm internal diameter (I.D.) Luna C18 (2) column with a particle size of 3 µm (Phenonemex, Aschaffenburg, Germany) at a flow rate of 0.5 mL min^{-1}. The mobile phases consisted of 0.05% phosphoric acid (A) and acetonitrile:0.2% formic acid (B) and were run at a 60-min linear gradient from 0 to 60% solvent B, followed by a 1-min linear gradient from 60 to 90% solvent B, plateauing for 4 min. The column was then flushed with 100% solvent A for 7 min.

P450 genes expression

To ascertain the involvement of *CYP72A219_4284* in the newly identified resistant population, a gene expression experiment was conducted. The aim was to assess the expression levels of *CYP72A219_4284* and determine its potential role in conferring resistance within the population. Plant growth and herbicide application at 91 g a.i. ha^{-1} were performed as mentioned before. Nebraska sensitive population [30] was used as a second negative control in this experiment. Four plants of each population were used for gene expression. Tissue of the first and second youngest leaves were harvested before and 6 h after herbicide application, respectively. Tissue was collected into 2 mL Eppendorf tubes and placed in liquid nitrogen. The tubes were kept in -80 ºC for further analysis. Tissue was ground with 3 mm stainless steel beads in a TissueLyser (Qiagen) with intensity of 30 for 1 min. RNA was isolated using Direct-zol RNA Miniprep from Zymo Research. cDNA synthesis was performed using ProtoScript® II First Strand cDNA synthesis kit using 1 µg of RNA and purified with DNase I. Gene expression analysis and conditions were the same as defined in the section of gene validation. The reference genes evaluated in the experiment was *18S rRNA*, whereas the candidate genes tested were *CYP72A219_4284* and *CYP81E8*.

Gene copy number

The six most resistant Palmer amaranth populations to tembotrione described above, along with NES, NER and another sensitive population IHX_3361 were used for *CYP72A219_4284* gene copy number analysis. Plants growth, tissue collection and DNA extraction were performed as described before. Genomic DNA at 50 ng/µL was used for relative gene copy quantification using a modified method $2^{-\Delta\Delta ct}$ [48]. The *ALS* gene was used as a low-copy control gene. The primers for *ALS* were used from previous research studying *EPSPS* copy number in *Amaranthus palmeri* [26]. Relative quantification of *CYP72A219_4285* was calculated with a modified method ΔCt (Ct, ALS – Ct, CYP). qPCR conditions were the same as described in the section for gene expression validation. Each population had eight biological samples run in two technical replicates.

Mapping of HPPD resistance in NER Palmer amaranth population

Dose-response and segregation analysis.

A dose-response was conducted to define delimiting rate to best differentiate the S × S (NES) and R × R (NER) populations. Seed of NES, NER and two F1 populations NERmale × NESfemale (crosses A and B) were sown on soil and then germinated in greenhouse at 28º C, photoperiod of 16h light and transplanted to 4 × 4 cm inserts after seven days. The dose response experiment consisted of 11 doses, which were 0, 1/128, 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 1, 2 and 4 × the label rate (91 g a.i. ha^{-1}). Each dose was applied on two plants with eight replicates. The herbicide tembotrione (Laudis, 419 g a.i. L^{-1}, Bayer, Leverkusen, Germany) was applied together with 1% v/v MSO using an automated spray chamber (Greenhouse Spray Chamber, model Generation IV) using a TJ8002E nozzle, calibrated to deliver 200 L ha^{-1} at a pressure of 280 kPa and speed of 1.2 m s^{-1}. Survival and fresh shoot mass were recorded 28 d after application and fitted to a three-parameter log-logistic model using the drc package in R [49].

Segregation in the pseudo-F2 population was performed in response to a delimiting rate of 77 g ha^{-1} with 606 plants in cross A and 1,277 plants in cross B. At 28 d after tembotrione application, plants were rated and shoot fresh mass was measured. For QTL analysis, the 71 and 110 most susceptible plants and 49 and 91 most resistant plants, from cross A and B, respectively, were selected based on survival and shoot fresh mass. Parental NES and NER were grown and submitted to the same herbicide screen and 20 of each population were selected for QTL mapping.

Library preparation and sequencing

Plant tissue was collected before herbicide application and had DNA extracted from single leaves following a modified CTAB method. DNA samples were assessed for quality using Qubit. Double-digest restriction site-associated DNA sequencing (ddRADseq) libraries were generated with ApeKI and sequenced using NovaSeq S4 with 150 bp paired-end reads (Illumina). The average yield was around 7.4 M reads per library and the mean quality scores were over Q30 for all libraries. The sequencing was performed at the University of Minnesota Genomics Center.

QTL identification and markers

The ddRAD-seq libraries were trimmed using trimmomatic v0.36 [50], aligned to the reference male genome of *Amaranthus palmeri* (scaffold file - v1.1, id55760) [37] with Burrows-Wheeler Aligner [51] and variants called using a GATK 4.2.0 [52]. Samples were computationally binned by shoot fresh mass. The variant sites were separated by SNPs and insertion/indels and hard filter was performed as following for SNPs - QD < 2.0, QUAL < 30.0, SOR > 4.0, FS > 20.0, MQ < 50.0 and insertion/indels QD < 2.0, QUAL < 30.0, FS > 200.0. The variant sites were filtered based on depth (at least five reads) and selected only variants that were homozygous. R/qtl2 package was used to find QTLs[53]. The analysis was performed separately

for cross A and B and combined A + B. QTLs intervals were calculated using Bayes credible intervals with the function bayes_int(). To determine the threshold for identifying potential QTL, 1000 permutation tests were conducted at a confidence level of 95%. The critical F value obtained from this analysis was used as the criterion to declare putative QTL. The functional analysis of the genes present in QTLs was conducted using DAVID Bioinformatics Resources v6.8[54].

Statistical analysis

The alignment of the RNA-sequencing experiment was used to analyzed differentially expressed genes (DEGs) using DEseq2 (v1.20.0)[55] package in R (v4.0.4)[56]. Populations S and R were considered as factor of conditions. Samples from cross A and cross B were pooled together for the analysis. Contrast analysis performed were R versus S before treatment (0 h), and 6 and 12 h HAT, and 6 h versus 0 h and 12 h versus 0 h for each population. The data were submitted for shrinkage log2 foldchange (LFC). Counts were normalized by creating a "virtual reference sample" using geometric mean of counts over all samples for each gene [57]. Principal component analysis (PCA), volcano and MA plots were performed on the normalized gene expression data. Genes with less than ten reads were excluded from analysis. A P-adjusted value <0.05 cutoff and log2 fold-change of 1 was used to identify DEGs.

The validation of candidate P450 genes in NER and NES was performed in four biological samples and two technical replications. *GUS* expression analysis was performed in three biological samples composed by six leaf disc and two technical replications. Analysis of the *CYP72A219_4284* and *CYP81E8* role in the newly resistant Palmer amaranth populations from 2019 were performed using four biological samples and two technical replications. The mean Ct values and the standard deviation were calculated by treatment. A melt curve analysis confirmed

the presence of a single amplified product for each reaction, based on presence of a single melting temperature consistent across samples for each gene. Relative transcript abundance was calculated using $2^{(-\Delta\Delta Ct)}$ method [48]. The reference population used for the P450 validation analysis was NES untreated. For P450 role analysis in the newly resistant populations, the sensitive control (IHX_3361) before herbicide application was used as control. The reference treatment used for the calculation of *GUS* expression was control, consisting of infiltration of *Agrobacterium* not carrying the pBI121 vector. Fisher's LSD's test ($p < 0.05$) was used to compare the relative expression between treatments for each experiment.

For the whole-plant dose responses, the statistical software R v.3.5.3[49] was used for data analysis. Data were adjusted using the three-parameter log-logistic model with the function modelFit () from the drc package[49]. The herbicide doses that caused a 50 % reduction in each variable were estimated using the model: Y=d/1+exp[b(logx-loge)], where *d* is the upper limit, *b* is the slope, *x* is the dose, and *e* is the dose that causes 50 % reduction in Y. The statistical difference between the resistant biotype and the sensitive biotype for ED_{50} (survival) or GR_{50} (fresh shoot weight) was calculated using the function EDcomp (). Resistance index (RI) was calculated using the ratio of GR_{50} values of each biotype with sensitive biotype. Graphs were generated using GraphPad Prism version 8.2.1 (San Diego, California).

For ^{14}C-tembotrione metabolism analysis, the identification of the main metabolites M1, M2, M3, M4, and M5 were based on previous results [30] and retention time. Peak area (% of recovered) was obtained from HPLC and used to quantify and compare metabolites. The means of six replications of each biotype were compared to IHX_3361 (sensitive) population by Dunnett's multiple comparison.

RESULTS

Differentially expressed genes (DEGs)

A total of 20,846 genes were analyzed for differential expression (DEGs) out of the initial set of 29,758 genes assigned by featureCounts. Genes with less than ten reads were excluded from the analysis. After applying data shrinkage, a gene-wise dispersion curve was fitted using the DESeq2 model (Supporting Information Figure 3-2). There was favorable dispersion pattern, with decreasing dispersion as the mean expression levels increased, indicating a good fit of the DESeq2 model to the analyzed data. The principal component analysis (PCA) plot revealed that PCA1 and PCA2 accounted for 35% and 21% of the variation in the data, respectively (Supporting Information Figure 3-3). PCA1 separated cross A from cross B, while PCA2 exhibited separation of gene clusters based on HAT within each cross. When clustering all 72 analyzed transcriptome samples, two major clusters emerged, with samples from different crosses grouped together, indicating similarities in gene expression response (Supporting Information Figure 3-4).

DEGs were observed in different comparisons. In the contrast comparison of R versus S at constitutively (0H), 6 H and 12 H after herbicide application, R had 37, 2, 9 up-regulated genes and 3, 2, 4 down-regulated genes, respectively (Figure 3-2A). Among these genes, only two genes were commonly up-regulated (Figure 3-2A). For the analysis of the genes that were responsiveness to herbicide application for each biotype, it was identified that 39, 530 and 106 genes were up-regulated in S and 6, 185 and 160 were up-regulated in R for the comparations of time 6 vs 0 HAT, 12 vs 0 HAT and 12 vs 6 HAT, respectively (Figure 3-2A). The heatmap illustrates the expression of DEGs from these comparisons and indicates a separation in clusters of resistant and susceptible plants (Figure 3-2B).

Several genes associated with the detoxification of xenobiotic substances were consistently identified, suggesting an enhanced genetic metabolism in plants with resistance. These genes included four cytochrome P450 genes, four glutathione-S-transferase genes, seven glycosyltransferase genes, a disease resistance protein, and a detoxification 27 gene (Supporting Information Table 3-3). Additionally, two MADS-box transcription factors of types 23 and 27 were found, indicating a potential regulation of gene expression in resistant plants. The resistance mechanism in the resistant population was characterized by an increased metabolism[30], and among all differentially expressed genes (DEGs), our focus was on the four cytochrome P450 genes that were consistently up regulated in the resistant plants. These genes were identified with the IDs MAKER_29886, MAKER_25717, MAKER_10107, and MAKER_25718, and exhibited fold changes of 18.3, 11.8, 6.2, and 5.4, respectively (Figure 3-2B-D and Supporting Information Table 3-3). None of these genes had differential expression at the 6 h time point compared to the susceptible plants (Figure 3-2C), and only one gene (MAKER_29886) exhibited a reactivation at the 12 h time point (Figure 3-2C). This indicates that the susceptible plants responded to the herbicide treatment (Figure 3-2C), where an increased transcription of these genes was observed at 6 and 12 HAT. However, the resistant plants consistently expressed these cytochrome P450 genes without any treatment.

Constitutively expressed P450s, namely MAKER_25717, MAKER_25718, and MAKER_29886, have been designated as *CYP72A219* due to their significant similarity to *CYP72A219* from *Spinacia oleracea* (NCBI: XM_056843192.1), with identities of 76.2%, 74.3%, and 76.3% respectively. Similarly, the gene ID MAKER_10107 has been assigned the name *CYP81E8* due to its high similarity (73.3%) to *CYP81E8* from *Chenopodium quinoa* (NCBI: XP_021724107.1). Hereafter, the genes MAKER_25717, MAKER_25718,

MAKER_29886, and MAKER_10107 will be referred to as *CYP72A219_4284*, *CYP72A219_4285*, *CYP72A219_7285*, and *CYP81E8* respectively. The number after CYP gene name indicates the gene number ID in the reference genome of Palmer amaranth.

SNPs in cytochrome P450 candidate

There were no consistent SNPs or amino acid changes between S or R in important domains of the P450s [58]. *CYP72A219*_4284 has a coding sequence of 1544 bp and protein with 517 amino acids. CYP72A219_4285 has a coding sequence of 1578 bp and 526 amino acids. Both populations share the same two alleles. *CYP72A219_7285* has a length of 1394 bp and 464 amino acids. Susceptible plants have two same alleles and R plants have a second allele with amino acid substitution of Thr307Ser, Phe388Leu, and Ile407Val. The former two *CYP72A219* (4284 and 4285) are localized in chromosome four and the latter is localized in chromosome 16. *CYP72A219_4284* has an identity with *CYP72A219_4285* and *CYP72A219_7285* of 76.1% and 55.1%, respectively. Two alleles of *CYP81E8* were present in S and R samples. The gene has a length of 1470 bp and 1479 bp in each allele. The short allele has a deletion of 18 bp in position 35, and the longest allele has a deletion of 9 bp in the position 37 when aligned to the draft genome (Supporting Information). In the available Palmer amaranth genome, this specific region of the gene exhibits a triplication of PPS amino acids, which is likely due to incorrect assembly. All cytochrome P450 possess a conserved cluster of proline in the membrane hinge. Similarly, the sequenced cytochrome P450 exhibits this conserved region. For more detailed gene sequences and alignments information, refer to the Supporting Information.

In the phylogenetic tree analysis, *CYP72A219* genes were cluster with the *CYP72A31* from *Oryza sativa*, which can metabolize bispyribac-sodium and bensulfuron-methyl [59]. *CYP81E8* was clustered with other P450 of family 81 from grasses but with some significant distances between

them (Figure 3-3). This gene has a low identity when aligned with the other P450s. *ApCYP81E8* has an identity of 40.5%, 41.0% and 39% with *ZmCYP81A9* (Genbank: EU955910.1), *LrCYP81A10v7* (Genbank: MK629521.1) and *EpCYP81A12* (Genbank:AB818460.1), known to metabolize herbicides in grasses [9, 16, 60, 61].

P450 qPCR validation

To confirm the upregulation of candidate cytochrome P450 genes in NER plants, their gene expression was measured before and after the application of tembotrione. The genes *CYP72A219_4284*, *CYP72A219_4285*, and *CYP72A219_7285* were upregulated in NER plants at 3, 6, and 12 HAT, respectively (Figure 3-4). Among the *CYP72A219* genes, *CYP72A219_4284* had a higher upregulation at 3 and 6 h after herbicide treatment. Additionally, *CYP81E8*, located on chromosome 4, had upregulation specifically at the 6 h time point following herbicide treatment (Figure 3-4). In contrast, *CYP72A219_4286*, which was not upregulated in the RNA-seq data, displayed similar results, indicating a lack of responsiveness to herbicides. These observations deviate from those obtained in the RNA-seq analysis, which indicated that these genes are constitutively expressed in resistant plants (Figure 3-2). However, this discrepancy arises due to the utilization of the parental populations NES and NER for gene validation, which exhibit significant variation in their response to tembotrione. Consequently, these results may differ from those identified in the RNA-seq analysis. Nevertheless, despite the disparity, the consistency of these findings indicates that the upregulation of these genes in NER plants differs from that in NES plants.

***CYP72A219* metabolizes tembotrione**

The yeast strains WAT11 and WAT21 were genetically transformed using the pYES2 vector containing the candidate cytochrome P450 genes. Following a 24 h incubation period with

tembotrione, some transformed yeast treatments converted tembotrione into hydroxy-tembotrione. The chromatogram of the empty vector pYES2 revealed solely the presence of the parent tembotrione peak at a retention time of 2.9 min (Figure 3-5). Notably, the *Nsf1* gene from corn exhibited a pronounced affinity for tembotrione, resulting in the production of a hydroxy-tembotrione peak at a retention time of 2.7 min and a dihydroxy tembotrione peak at 2.5 min (Figure 3-5). Among all the candidate P450 genes tested, only *CYP72A219_4284* demonstrated the ability to convert the parent tembotrione into hydroxy-tembotrione. Both alleles of this gene in either the WAT11 or WAT21 strain were capable of metabolizing the herbicide. Conversely, *CYP72A219_4285*, *CYP72A219_7285*, and *CYP81E8* did not exhibit any ability to metabolize tembotrione. Based on these results, it can be concluded that only *CYP72A219_4284* possesses the capacity to recognize the tembotrione molecule and effectively metabolize it.

CYPs promoter analysis

The DNA segment amplified upstream *CYP81E8* was 1,250 base pairs for both the sensitive and resistant plants (Figure 3-6). For the promoter of the *CYP72A219_4284* gene on chromosome four (MAKER-25717), the primers amplified regions ranging from 1991 to 2004 base pairs for S plants and from 1747 to 1768 base pairs for R plants. The needle alignment between two samples revealed 100% identity between the promoters of S and R plants for *CYP81E8* promoter (Supporting Information Figure 3-5), indicating there was no difference between S and R plants, except for one S sample that exhibited a higher similarity to the genome sequence (Figure 3-6A). In contrast, needle alignment of *CYP72A219_4284* promoter revealed sequence differences between the samples S and R, with identity of 76.8% (Supporting Information Figure 3-5). Resistant plants had unique insertions, such as a cytosine at position 45 base pairs, an eight-base-pair insertion at position 458 base pairs, an eight-base-pair insertion at

position 1247 base pairs upstream, and a 31-base-pair insertion at position 1291 base pairs upstream. Sensitive plants, on the other hand, had unique insertions, including a 13-base-pair insertion at position 82, a 10-base-pair insertion at position 309 base pairs, and a 277-base-pair insertion at position 746 base pairs upstream of the gene (Figure 3-6B and Supporting Information Figure 3-5). The 277-base-pair insertion in sensitive plants was further analyzed by blasting against the Palmer amaranth genome, resulting in hits on chromosomes 16, 4, and 9 were found, which indicates a genetic variability within the sensitive population.

A total of fifteen significant binding-site motifs were identified in the promoter region of *CYP81E8*. Interestingly, all of these motifs were found in both sensitive (S) and resistant (R) promoters (Supporting Information Figure 3-6). In the case of the *CYP72A219_4284* promoter, a total of twenty-five motifs were discovered. Among these, two motifs were specific to the S promoter, two were specific to the R promoter, and twenty-one motifs were shared between S and R (Figure 3-7 and Supporting Information Figure 3-6). Notably, the R promoter exhibited replication of four motifs. The specific function of the R motif (MAAGYMKMATTTACM) in the *CYP72A219_4284* promoter was not determined using GOMo analysis. However, the motif GTTRRTKWGKCTTKWWRKCWTKMWGATWWWTACTRYTAWTAYMWYTSTCT was identified as a binding site for transcription factors such as MYB88 (AT2G02820), suppressor factor AIF1 (AT3G05800), and abiotic stress-responsive transcription factor DREB (AT1G77200) based on GOMo analysis. Additionally, the motifs TCWYYMTTTTTTTCTCCATCATACAAAATAAATGAGACTGW, which is a binding site for zinc finger transcription factors like C2H2-type (AT2G28710), and CTAATRTCTYATTTTYAAATG, which is a binding site for NAC transcription factor NAC050 (AT3G10480), were duplicated in the R promoter (Figure 3-7). Furthermore, the motif

GTTATTTAGTAACTTWKYGTD, which serves as a binding site for the MYB gene transcription factor, was triplicated in the R promoter (Figure 3-7). Based on these findings, it appears that the high expression of *CYP81E8* in HPPD-resistant plants may be regulated by a trans-regulatory element. Conversely, the constitutive increased expression of *CYP72A219_4284* may be regulated by a *cis*-element due to differences in the promoter sequence that co-segregate with resistance in the pseudo-F2 generation, however, a *trans*-element is not discarded.

Transient GUS expression and histochemical analysis

GUS expression was analyzed 96 h after Agroinfiltration, where the vector pBI121 carrying different promoters with the *GUS* gene was used. The *35S::GUS* construct, serving as the positive control, exhibited a robust upregulation of the *GUS* gene. Similarly, the *750_R::GUS* and *1500_R::GUS* promoters had the highest average *GUS* expression compared to other treatments. These two treatments did not differ significantly from the *35S::GUS* control; however, due to the variability in the data, they also did not differ significantly from the negative control *chunk::GUS* (Figure 3-8A). While results were not statistically significant, the promoters derived from resistant plants with lengths of 750 bp and 1,500 bp demonstrated a trend of increased *GUS* expression. These findings were further supported by the GUS histochemical assay, which revealed a stronger intense blue reaction in these two treatments (Figure 3-8B).

Whole-plant dose response, metabolism and *CYP* expression in different herbicide resistant Palmer amaranth

Based on the whole-plant dose response (Figure 3-9), the confirmation of Palmer amaranth populations with suspected tembotrione resistance revealed distinct variations. The sensitive population had an LD_{50} and GR_{50} parameters of 22.8 g and 7.3 g, respectively, and the suspected resistant populations had a significantly higher LD_{50} values ranging from 55.4 g (WR2019-141)

to 183 g (WR2019-137) (Figure 3-9C). Similarly, for GR_{50}, the resistant populations ranged from 19.8 g (WR2019-199) to 50.1 g (USA 12001) (Figure 3-9D). In general, the resistance index ranged from 2.4 (WR2019-144) to 7.9 (WR2019-044) for LD_{50} and from 2.7 (WR2019-199) to 6.8 (WR2019-137) for GR_{50} (Figure 3-9C and D). The findings conclusively establish resistance to tembotrione in all populations examined, demonstrating varying levels of resistance across the tested samples. The logistic parameter tables are indicated in the Supporting Information Table 3-4 and 3-5.

Tembotrione metabolism was assessed in the Palmer amaranth populations using ^{14}C herbicide (Figure 3-10). The hypothesis was that enhanced metabolism in the field-collected populations is the main mechanism of resistance, similar to the previously studied tembotrione-resistant population NER [30]. The reverse-phase HPLC method was able to detect the five main metabolites and parental tembotrione. The method used in the present study is slightly different from the previous study of tembotrione metabolism [30], hence, retention times are slightly different. Parental tembotrione had a retention time of 66.9 min and its major metabolites had retention times of 47.3 (M1), 50.1 (M2), 56.3 (M3), 59.0 (M4), and 67.7 min (M5) (Supporting Information Figure 3-7 and 3-8).

In the previous study, the metabolic pathway of HPPD in herbicide-resistant plants was thoroughly examined and discussed (Supporting Information Figure 3-7). NER population had enhanced metabolism of tembotrione to hydroxy-tembotrione at a faster rate than NES, occurring as early as 6 HAT [30]. In our study, there was no difference in the metabolites peak areas at 6 HAT; however, NER had a smaller peak of parental ^{14}C-tembotrione than sensitive plants at this time point. When 90% of the parental ^{14}C-tembotrione was detected in sensitive plants, in NER it was detected only 59% of total recovered ^{14}C-tembotrione. The same results were found in

WR2019-140, -137, -199, -200, -044, -198 and -273 populations, which showed smaller area peak of parental ^{14}C-tembotrione at 6 HAT than sensitive plants, with a range of 52 to 67 % of total recovered.

Analyzing the hydroxy-tembotrione (M3/M4), the main metabolite which confer herbicide resistance, NER had higher % area peak than sensitive population at 12 HAT. The other resistant populations were similar, except for WR2019-044 and WR2019-273. NER and WR2019-034, -274, -141, -140, -137, -199, -200 had a higher hydroxy-tembotrione peak than sensitive plants at 24 h and along with WR2019-044 at 48 h after herbicide application (Figure 3-10D). M3 and M4 despite having two separate peaks, were identified to have the same mass, suggesting they are isomers with the hydroxy group attached in different locations of the tembotrione[30] (Supporting Information Figure 3-7). The faster hydroxy-tembotrione production in the majority of the populations suggests that enhanced metabolism capacity is the main mechanism of resistance of these populations (Supporting Information Figure 3-7).

The production of M1 was higher in some population at 24 HAT than sensitive and NER. The production of M2 was not detected in the majority of the populations. The production of M5 was lower or higher in different populations when compared to sensitive plants. Altogether, the different metabolites production among the populations, suggests a secondary different pathway through tembotrione can be metabolized. Different resistance mechanisms than enhanced metabolism could be associated in the populations WR2019-273. This conclusion arises from the observation that this population had minimal hydroxy-tembotrione production across various time points (Figure 3-10). However, this observation can be attributed to the lower level of resistance detected in this population (Figure 3-9).

With the hypothesis that *CYP72A219_4284* gene is involved in the metabolism of tembotrione in the newly characterized tembotrione-resistant populations collected from the field in 2019, gene expression analysis was performed. *CYP72A219_4284* was upregulated in majority of the resistant populations, either before or 6 h after herbicide application (Figure 3-11). The populations WR2019-034, -137, -140, -199, and -274 had a significantly higher relative expression compared to the sensitive population of 363.6, 350.8, 539.9, 335.0 and 657.8. Other populations, including NER, WR2019-141, -144, and -198, had over 100-fold up-regulation of *CYP72A219_4284* compared to the sensitive population; however, due to data variability within each population the data was not significant, indicating potential segregation. At 6 h after herbicide application, expression of *CYP72A219_4284* responded even in sensitive plants (292.7), but this gene had higher expression at this time point in the resistant populations WR2019-034, -137, -140, -199, and -274. The newly identified tembotrione-resistant populations collected in 2019 demonstrated constitutive and responsiveness overexpression of the *CYP72A219_4284* gene, with some populations even exhibiting higher expression levels than NER. In contrast, the expression of *CYP81E8* was not constitutively or responsiveness to tembotrione treatment in Palmer amaranth populations, except in WR2019-140 (Figure 3-11). The gene copy number of *CYP72A219_4284* remained consistent across all resistant populations when compared to sensitive plants (Supporting Information Figure 3-9). Therefore, the increased expression of this *CYP72A219_4284* is not due to increased gene copy number. The higher gene expression of *CYP72A219_4284*, constitutively and after tembotrione application in resistant plants, suggests its crucial role in herbicide resistance in Palmer amaranth, potentially attributed to enhanced metabolism.

Degree of dominance and QTL identification

Increasing doses of tembotrione were applied to parental NER and NES plants together with two F1 populations generated from two crosses. The sensitive NES population had an ED_{50} of 7.9 g and NER had an ED_{50} of 37.3 g (Table 3-1). The ED_{50} found in the present study was smaller than the one previously reported for these populations [30]; however, the resistance index (RI) between these two populations was similar. In the present dose-response the RI was 4 compared with 3.3 from the previous study [30]. Both F1 populations were intermediate between NES and NER in response to tembotrione (Figure 3-12), having ED_{50} of 16.6 and 23.6 g, respectively, for cross A and B (Table 3-1). The degree of dominance was calculated and indicated that HPPD resistance in NER is a co-dominant or semi-dominant trait (cross A = 0.42 and cross B = 0.67) based on the formula proposed by Bourguet and Raymond (1998)[62]. When tembotrione at a rate of 77 g ha^{-1} was applied in the pseudo-F2 populations resulting from the cross between A and B, they exhibited survival rates of 11.4% and 8.8%, respectively. These survival rates are similar to the parental homozygous line NER, which was cultivated alongside the pseudo-F2 plants and had a survival rate of 12.4%.

Frequency distribution for QTL mapping on pseudo-F2 plants had two distinct peaks for each cross, with one peak each indicating phenotypic similarity with each parental population. A normal distribution of R plants had higher biomass (red bars) similar to the average biomass of parental resistant NER (RxR) population indicated by the dashed red line (Figure 3-13A). The same was identified for S samples, having a normal distribution around the average biomass of the parental NES (SxS) plants (Figure 3-13A).

The variant calling analysis identified over 3.4 million single nucleotide polymorphism (SNP) sites in the Palmer amaranth genome. After filtering for homozygosity at each site and a minimum read depth of five, a total of 4,404 SNP variants were detected throughout the entire

Palmer amaranth genome. Genome scanning in cross A did not detect any QTLs surpassing the LOD threshold of 6.4 after conducting a permutation test. However, when examining pseudo-F2 plants from cross B, significant QTLs with high effects were found in multiple scaffolds overpassing LOD threshold of 6.7. The scaffolds include scaffold 10 with LOD of 11.6, scaffold 81 with LOD of 10.1, scaffold 6 with LOD of 9.5, and scaffold 14 with LOD of 8. Each QTL are accompanied by other significant QTLs nearby (Figure 3-13B, Supporting Information Table 3-6). For instance, in Scaffold 10, there are three QTL peaks, the biggest peak with a LOD score of 11.6 at position 11,877,263, and two others peaks with LOD scores of 8.4 and 7.9 at positions 10,336,852 and 10,337,656, respectively (Supporting Information Table 3-6). The same happened in scaffold 81, 6 and 14 with nearby peaks (Supporting Information Table 3-6). Combining the datasets from cross A and cross B increased the sample size and enhanced the statistical power for the analysis. As a result, the genome scan revealed a more robust QTL effect on scaffold 10, precisely at the same genomic location 11,877,263. The QTL exhibited a higher LOD score of 12.2, indicating a stronger and more significant association with the target trait (Supporting Information Table 3-6). Additional QTLs were found on scaffolds 6, and 14, but with lower effects.

The proportion of phenotypic variation explained by each identified QTL was calculated. For cross B, the QTL located in scaffold_10 accounted for 23% of the phenotypic variation observed in the data. Additionally, the QTLs in scaffold 81, 6, and 14 explained 20%, 19%, and 16% of the phenotypic variation, respectively. However, when the data from both cross A and cross B were combined, the overall phenotypic variation explained by the QTLs decreased. This decrease can be attributed to the absence of any QTL in cross A. Specifically, for the combined

data, the QTL in scaffold_10 explained 15% of the phenotypic variation, while the QTLs in scaffold 6 and 14 accounted for 10% each (Supporting Information Table 3-6).

In order to conduct further analysis, the combined dataset was chosen for functional annotation due to its larger number of genes and sample size. Additionally, a significant overlap was observed between the genes in the combined dataset and those identified in cross B. Therefore, the combined dataset provides a more comprehensive representation of the genes of interest and is thus suitable for functional annotation analysis. The QTLs found in scaffolds 10, 6, and 14 are localized on chromosomes four, eight, and fifteen of Palmer amaranth, respectively, and encompass 82, 196, and 25 genes, respectively (Supplemental Information Table 3-6). Functional annotation of these genes indicates a predominant molecular function of protein binding (GO:0005515), RNA binding (GO:0003723), and mRNA binding (GO:0003729), consistent with proteins involved in regulating other proteins or molecules through selective protein interactions [63] (Figure 3-14). The important TFs that were identified includes *DREB2A*, *WRKY* and GATA16 within the QTLs. The QTLs also include nine zinc finger proteins, four ubiquitin proteases, three F-box proteins, three ABC transporters, and three cytochrome P450 enzymes. A *CYP86A1* is localized in one of the QTLs in scaffold 6, *CYP707* and *CYP71A* are localized in the QTL of scaffold 14, which have small phenotype effects.

For each QTL, multiple markers were identified. In the case of the QTL in scaffold 10, a total of 54 markers were found, while for the QTL in scaffold 6, 94 markers were identified. However, for the QTL in scaffold 14, only a single marker was detected within a 1 Mb range upstream and downstream of the QTL region. The markers associated with scaffolds 10 and 6 displayed a consistent pattern with the evaluated phenotype, where the RR genotype exhibited higher biomasses, while the SS genotype had lower biomasses (Figure 3-15). In contrast, the

single marker found in scaffold 14 did not demonstrate the same consistency as the markers identified in the other QTLs (Figure 3-15I).

DISCUSSION

Metabolic herbicide resistance mechanisms in weeds represents a significant challenge in weed control practices[4]. These mechanisms enable weeds to metabolize a diverse array of herbicides from various chemical families. Identifying the genes responsible for metabolic resistance and comprehending their regulation holds immense potential for developing improved herbicides. In our current investigation, four P450 genes that are consistently upregulated were identified in HPPD-resistant Palmer amaranth, a troublesome weed species. These four genes did not have any polymorphisms in the coding sequences between R and S plants. Among these genes, only one, named *CYP72A219_4284*, deactivated the herbicide tembotrione when tested in a heterologous system. Furthermore, our study revealed that this specific gene is upregulated in other HPPD-resistant Palmer amaranth populations collected from various fields across US, supporting its involvement in conferring herbicide resistance to this globally significant species.

In the past, metabolic herbicide resistance in dicot weeds was relatively uncommon, with most documented cases of metabolic resistance and identified P450 genes occurring in monocots. However, recent studies have reported instances of metabolic resistance in dicots such as *A. tuberculatus* in 2017 [64], *A. palmeri* [30] in 2018, and *Raphanus raphanistrum* in 2020 [65]. These weed species exhibited resistance to HPPD inhibitors. Furthermore, in 2020, metabolic resistance to 2,4-D was identified in *A. tuberculatus* [66]. The identification of P450 genes capable of metabolizing herbicides in dicots is relatively limited. For example, *DsCYP77B34* from *Descurainia sophia*, a broadleaf weed found in Asian wheat fields, conferred resistance to ALS-, PPO-, VLCFA-, and PSII-inhibitors when expressed in transformed *Arabidopsis thaliana*[17].

Arabidopsis expressing *Raphanus raphanistrum* genes *RrCYP704C1* or *RrCYP709B1* was resistant to HPPD inhibitors such as mesotrione *(RrCYP704C1)*, tembotrione *(RrCYP709B1)*, and isoxaflutole *(RrCYP709B1)*[20]. In *Amaranthus palmeri*, only one study has identified *CYP81E8* as the primary gene responsible for metabolic resistance in a HPPD-resistant population from Kansas, USA (KCTR)[67]. This study found that the population exhibited enhanced metabolism of tembotrione, and *CYP81E8* expression, measured only by qPCR, was implicated as the likely cause of tembotrione resistance. However, our study indicates that although *CYP81E8* is upregulated, it is not the causal gene for tembotrione resistance in the NER population. Instead, we have identified *CYP72A219_4284* was identified as the gene capable of converting tembotrione to hydroxy-tembotrione. This finding represents the first study which validated the gene responsible for herbicide metabolism in Palmer amaranth.

CYP81E8 has been identified as the responsible gene for 2,4-D resistance in Amaranthus tuberculatus[68]. In a previous study, it was observed that this specific cytochrome P450 gene co-segregated with 2,4-D resistance based on RNA-seq analysis. However, it did not co-segregate with HPPD resistance, indicating that the genetic association was specific to 2,4-D resistance, but further gene validation was not performed[68]. Interestingly, *CYP81E8* in waterhemp and Palmer amaranth exhibits a high similarity, with a sequence identity greater than 92% (Supporting Information). A similar behavior of *CYP81E8* was observed in this study, with higher expression levels in resistant plants of Palmer amaranth; however, when tested in a heterologous system, this gene did not inactivate tembotrione. Based on this finding, we hypothesized that the NER population might have resistance to 2,4-D due to up-regulation of *CYP81E8*. The NER population had a resistance index ranging from 2 to 2.5, but complete control was achieved when 500 a.e. g ha^{-1} of 2,4-D was applied (data not shown). This suggests

that the NER population may indeed be evolving resistance to 2,4-D, however, resistance to tembotrione is not attributed to *CYP81E8* but rather to *CYP72A219_4284*.

Previous research has highlighted *CYP72A219* as a gene involved in herbicide resistance in weeds. Through RNA-seq analysis, a study indicated that among other metabolic genes, *CYP72A219* had constitutive expression in glufosinate-resistant *Amaranthus palmeri* plants, with eight times higher expression levels in resistant plants compared to susceptible ones, but no gene validation was performed[69]. Furthermore, *BsCYP72A219*, along with *BsCYP81Q32* from *Beckmannia syzigachne*, was upregulated in mesosulfuron-methyl-resistant plants following RNA-seq analysis[70]. These genes were validated in transgenic rice, where only *CYP81Q32* conferred resistance to ALS-inhibitors. Notably, upregulation of two genes and only one gene being actually responsible for the herbicide resistance, resembled the observed pattern in HPPD-resistant Palmer amaranth in the present study. Additionally, an alignment of *CYP81E8* from Palmer amaranth revealed a 92.7% identity with *CYP81Q32-like* from *Amaranthus tricolor* (XP_057543510.1), a recently sequenced *Amaranthus* species[71]. This finding indicates a very close relationship between *CYP81E8* and *CYP81Q32* and may even represent the same gene, designated with different names. The pattern of co-segregation and upregulation of these genes in different weed species suggests the presence of common *cis* or *trans*-elements that potentially regulate CYP expression in resistant plants.

Distinct promoter sequences for *CYP72A219_4284* were found in resistant pseudo-F2 plants compared to susceptible, while no such variations were observed for *CYP81E8*. The *GUS* expression analysis revealed a significant trend of upregulation when the promoter derived from resistant plants was inserted in the vector, specifically when the promoter region located 750 base pairs or more upstream of the gene was utilized. Our findings indicate that the promoter

segment positioned at least 250 base pairs upstream of the gene exhibited duplicated cis-elements and motif sequences, suggesting the possibility of enhancing gene expression. QTLs with high effects on the herbicide resistance trait, localized in chromosomes four, eight and fifteen were also identified. These QTLs include three cytochrome P450 genes and ABC transporters. While these are important genes and warrant further analysis, the gene responsible for metabolizing tembotrione was already identified through RNA-seq analysis and validated, suggesting that the new P450s in the QTL mapping might have a small effect or even not be crucial for the resistance mechanism of the NER population. The key findings by the QTL are the diverse set of regulatory genes, including key transcription factors that could play a crucial role in the upregulation of *CYP72A219* in resistant plants. Transcription factors such as *DREB2A*, *WRKY,* and *GATA16*, along with zinc finger family found in the mapping analysis are known for their involvement in abiotic stress responses [72-74]. Significantly, it should be highlighted that the identified QTL on chromosome four is located two Mbp upstream of *CYP72A219_4284* gene. This proximity suggests that the transcription factor (TF) within the QTL region may regulate the gene, potentially influencing its expression. This regulatory effect could be particularly relevant due to the presence of a distinct promoter sequence observed in resistant plants.

Promoters of P450s were also investigated in other weeds species. Nucleotide polymorphisms were observed in the promoters of *CYP81A12* and *CYP81A21* in the R alleles of multiple herbicide resistant *Echinochloa phyllopogon* plants; however, these single nucleotide polymorphisms did not exhibit a significant association with herbicide resistance[61]. Three SNPs were identified in the promoter region of *BsCYP81Q32* between the R and S variants, spanning approximately 567 to 1207 base pairs upstream of the gene. These SNPs were responsible for

increased expression of *GUS* in tobacco leaves transformed with the gene, suggesting the presence of a crucial sequence within that specific region[70]. Additionally, the researchers identified a transcription factor, *BsTGAL6*, belonging to the bZIP family, which demonstrated efficient binding to the promoter of *BsCYP81Q32*[70]. This finding suggests the involvement of a *trans*-element that regulates the expression of *BsCYP81Q32* in resistant plants.

To determine the responsible transcription factors involved in *CYP72A219_4284* regulation in the present study, further analysis is required. One approach is to conduct a ratiometric luciferase assay using candidate TFs and the promoter of *CYP72A219*[75]. This assay involves the detection of firefly luciferase and renilla luciferase, which can indicate the interaction between TFs and the resistant promoter of *CYP72A219*. Another potential method is the Yeast one-Hybrid assay (Y1H)[76]. This assay utilizes a reporter construct with the DNA of interest (promoter) upstream of a reporter gene. Additionally, a second expression vector is used to generate a fusion between the candidate TF of interest and the transcription activation domain (AD). If the TF binds to the promoter, AD will induce the reporter gene, allowing for detection and analysis. These approaches can provide valuable insights into the interaction between TFs and the resistant promoter of *CYP72A219*, aiding in the validation of gene regulation mechanisms.

In conclusion, our study successfully identified and validated the metabolic resistant gene, *CYP72A219_4284*, in the HPPD-resistant *Amaranthus palmeri* population NER. This gene demonstrated the ability to metabolize tembotrione into the main metabolite hydroxy-tembotrione. The identification of this gene as highly upregulated in various Palmer amaranth populations across the United States poses a potential threat to weed management strategies

targeting this species. We also discovered important QTLs associated with the resistance trait, warranting further investigation and validation.

To our knowledge, this is the first study to identify and validate a gene responsible for tembotrione resistance in Palmer amaranth. Future research focusing on validating the identified transcription factors will shed light on the regulation of P450 genes and enhance our understanding of their regulation in weed species.

TABLES

Table 3-1. Logistic equation parameters and resistance index (RI) for survival (% of untreated control) in parental *Amaranthus palmeri* herbicide-susceptible (NES), herbicide-resistant (NER), F1 cross A, and F1 cross B populations in response to tembotrione.

Population	b	ED_{50}				RI	p-value
		Dose (g ha^{-1})	Lower CI (95%)	Upper CI (95%)	SE		
NES	2.4	9.3	7.9	10.6	0.7	-	
NER	2.1	37.3	31.2	43.4	3.1	4	<0.001
F1 cross A	2.4	16.6	14.2	19.1	1.2	1.8	0.2
F1 cross B	2.6	23.6	20.15	27	1.7	2.5	<0.001

b: slope; ED_{50}: herbicide dose that causes a 50 % reduction in plant survival; CI: confidence interval of parameter ED_{50} ($\alpha = 0.05$); RI: resistance index = ED_{50} ratio between populations with NES.

FIGURES

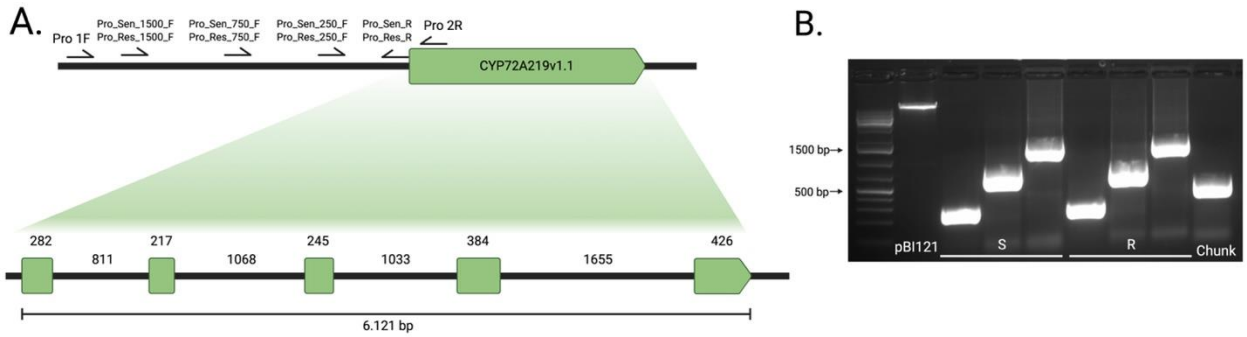


Figure 3-1. Illustration of *CYP72A219_4284* promoter, gene structure and agarose gel of promoter pieces amplification. A) Scheme representing *CYP72A219_4284* promoter with the primers localization to amplify different promoter sizes. B) Agarose gel from the digestion reaction of the different promoter sizes.

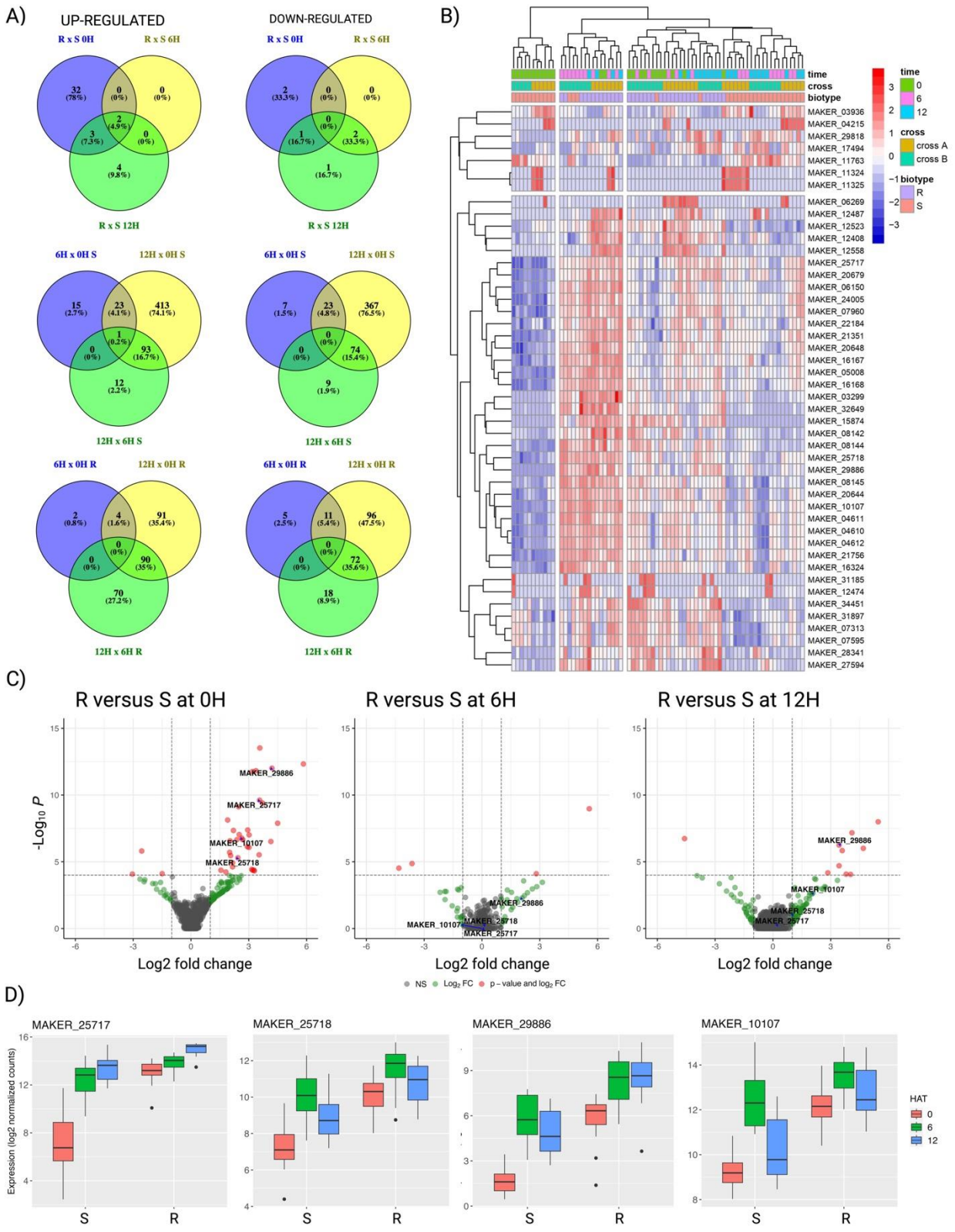


Figure 3-2. Analysis of differentially expressed genes (DEGs) in Pseudo-F2 *Amaranthus palmeri* susceptible (S) and HPPD-resistant (R). A) The Venn diagram of the up- and down-regulated

DEGs indicating unique and common DEGs for 18 different comparisons. B) Heatmap of top DEGs for the comparisons R vs S at 0H, 6 and 12 HAT. C) Volcano plots displaying gene expression differences between *Amaranthus palmeri* and treatments of contrast of R versus S before herbicide application (0H), contrast of R versus S at 6 h after herbicide application and contrast of R versus S at 12 h after herbicide application. Differentially expressed genes with significant thresholds of p-value <0.0001, adjusted p-value <0.05, log2 fold-change >1 or < -1 are represented by red circles. Dashed lines represent significance thresholds of adjusted p-value<0.05 and log2 fold-change >1 or < -1. D) The log2 normalized counts of top four cytochrome P450 genes in response to tembotrione in S and R population. Differently expressed genes had a threshold of adjusted p-value of 0.05 and log2 fold-change 1.

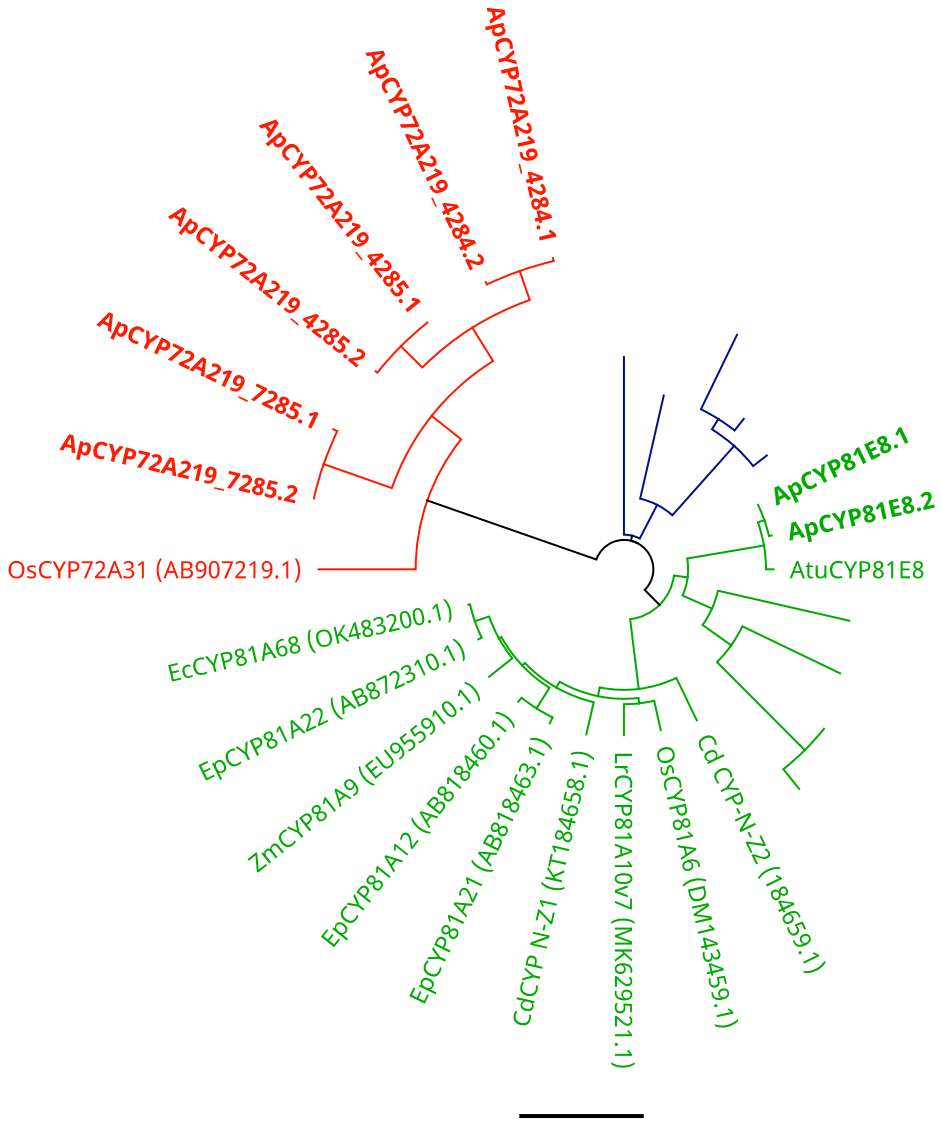


Figure 3-3. Phylogenetic analysis of CYP72A219_4284, CYP72A219_4285, CYP72A219_7285 and CYP81E8 proteins from *Amaranthus palmeri* and other P450s that metabolize herbicides from different species. Ap – *Amaranthus palmeri*, Atu – *Amaranthus tuberculatus*, At – *Arabidopsis thaliana,* Cd – *Cynodon dactylon*, Ec – *Echinochloa crus-galli*, Ep – *Echinochloa phyllogon*, Gm – *Glycine max*, Ht – *Helianthus tuberorus*, Lr – *Lolium rigidum*, Nt – *Nicotiana tabacum*, Os – *Oryza sativa*, Zm – *Zea mays.*

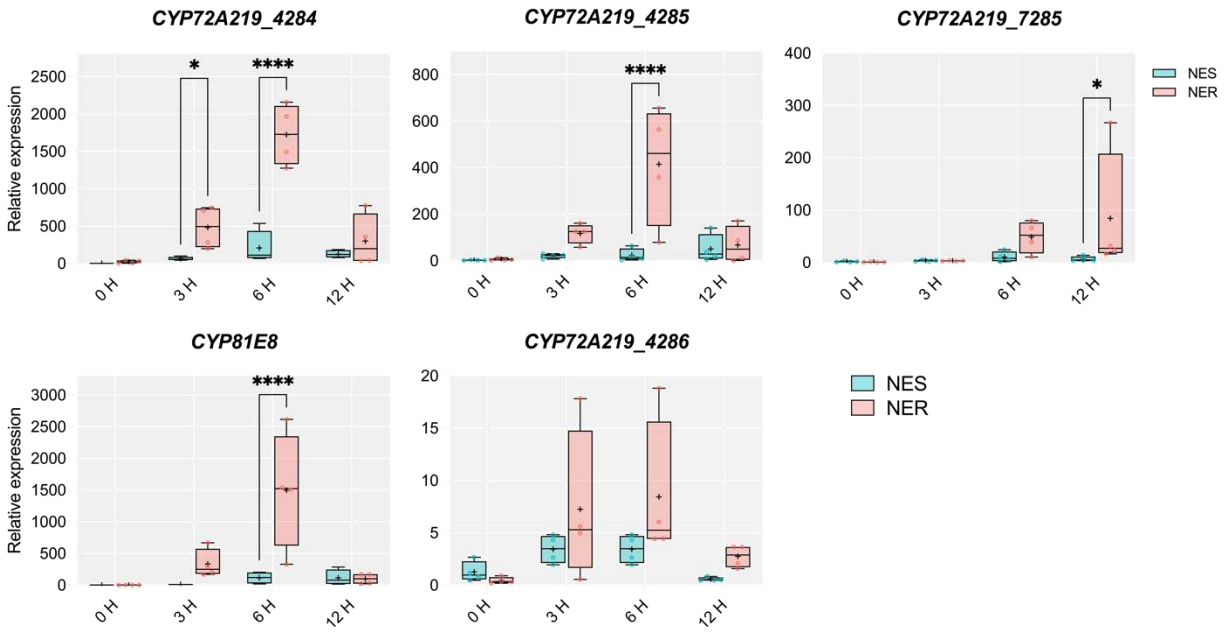


Figure 3-4. Relative gene expression of candidate cytochrome P450 genes in parental NES and NER plants. Boxplot indicates the average, the minimum and maximum values obtained from four biological and two technical replicates. Plants samples were collected before, 3, 6 and 12 h after tembotrione application (91 g a.i. ha^{-1}). The relative expression values were calculated by normalizing to the gene expression of NES untreated, using the average of the normalization gene *18S* and *Actin7*. Statistical differences at each time point were determined using the Fisher's LSD test, with asterisks indicating the significance level: * $p < 0.05$, and **** $p < 0.0001$.

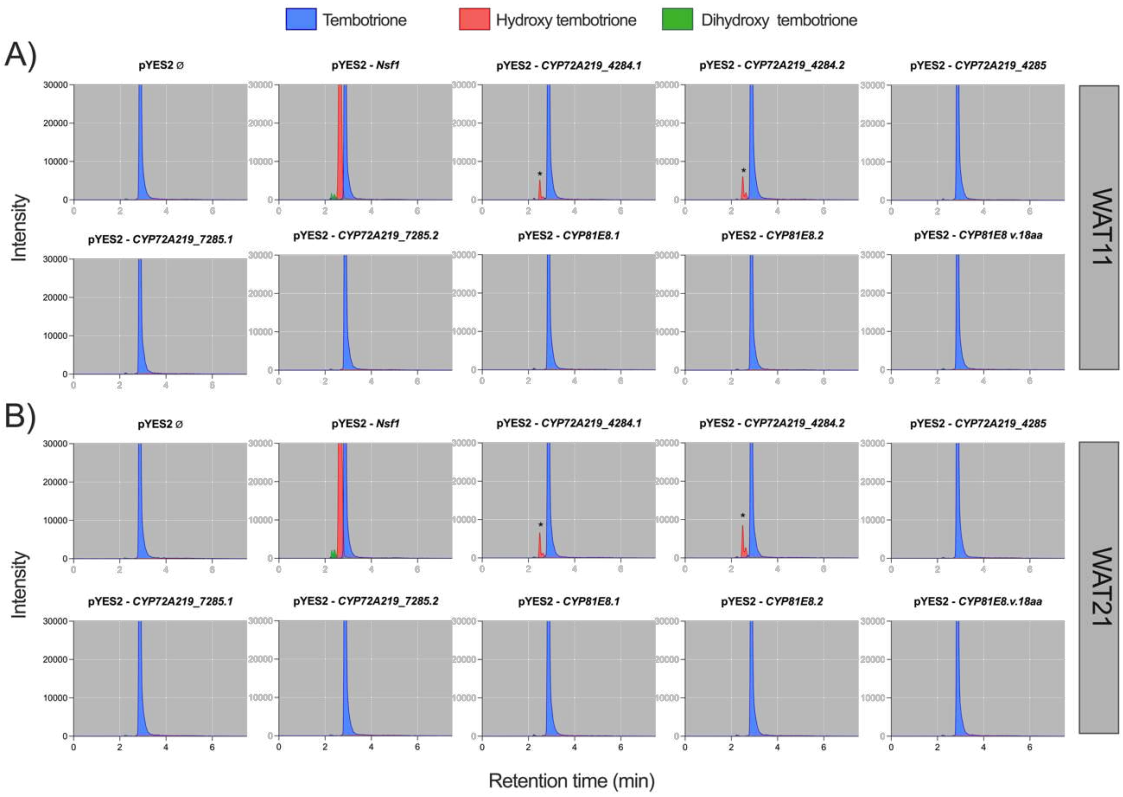


Figure 3-5. LC-MS/MS chromatogram of metabolites formed in yeast transformed with P450 genes from Palmer amaranth. Blue peak – parental tembotrione, read peak – hydroxy-tembotrione, green peak, dihydroxy-tembotrione. A) WAT11 and B) WAT21 yeast strain carrying *Arabidopsis thaliana* cytochrome P450 reductase 1 and 2, respectively. * at least 10-fold higher peak than the background signal.

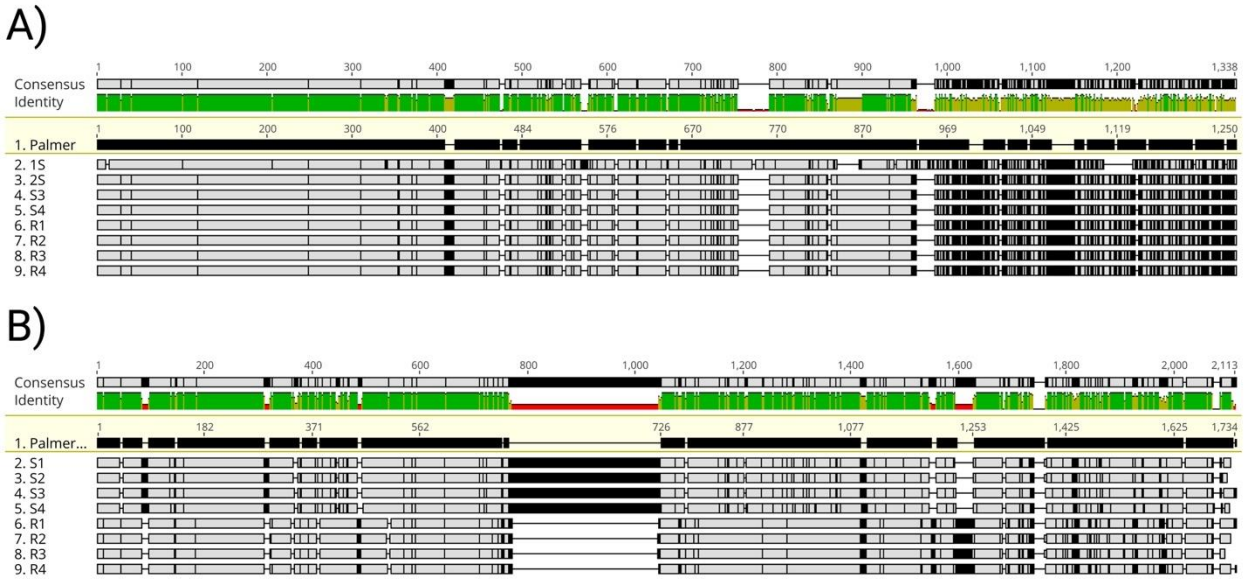


Figure 3-6. MUSCLE v5 multiple alignment of *CYP81E8* (A) and *CYP72A219_4284* (B) promoter sequences. The promoter sequences were reverse-oriented to facilitate the positions understanding. Black space indicates differences from the reference Palmer amaranth genome (id 55760).

Motif	Logo	Presence in promoter	Molecular Function
TCWYYMTTTTTTTCTCCATCATACAAAATAAATGAGACTGW		1x S 2x R	TF ZI
TATAATAHMTTATATTTRRGRTTATAAHAWADTATWWTBG		3x S	TF
GTRRRTKWGKCTTKWWRKCWTKMWGATWWWTACTRYTA WTAYMWYTSTCT		1x R	TF
GTTATTTAGTAACTTWKYGTD		1x S 3x R	TF
CTAATRTCTYATTTTYAAATG		1x S 2x R	TF
CTTTTCCATCYCCGTTTTTAAGCGGAATA		1x S	TF
MAAGYMKMATTTACM		2x R	-

Figure 3-7. Motif sequences found in *Amaranthus palmeri CYP72A219_4284* promoter in sensitive and resistant pseudo-F2 plants. TF, transcription factor activity. ZI, zinc ion binding.

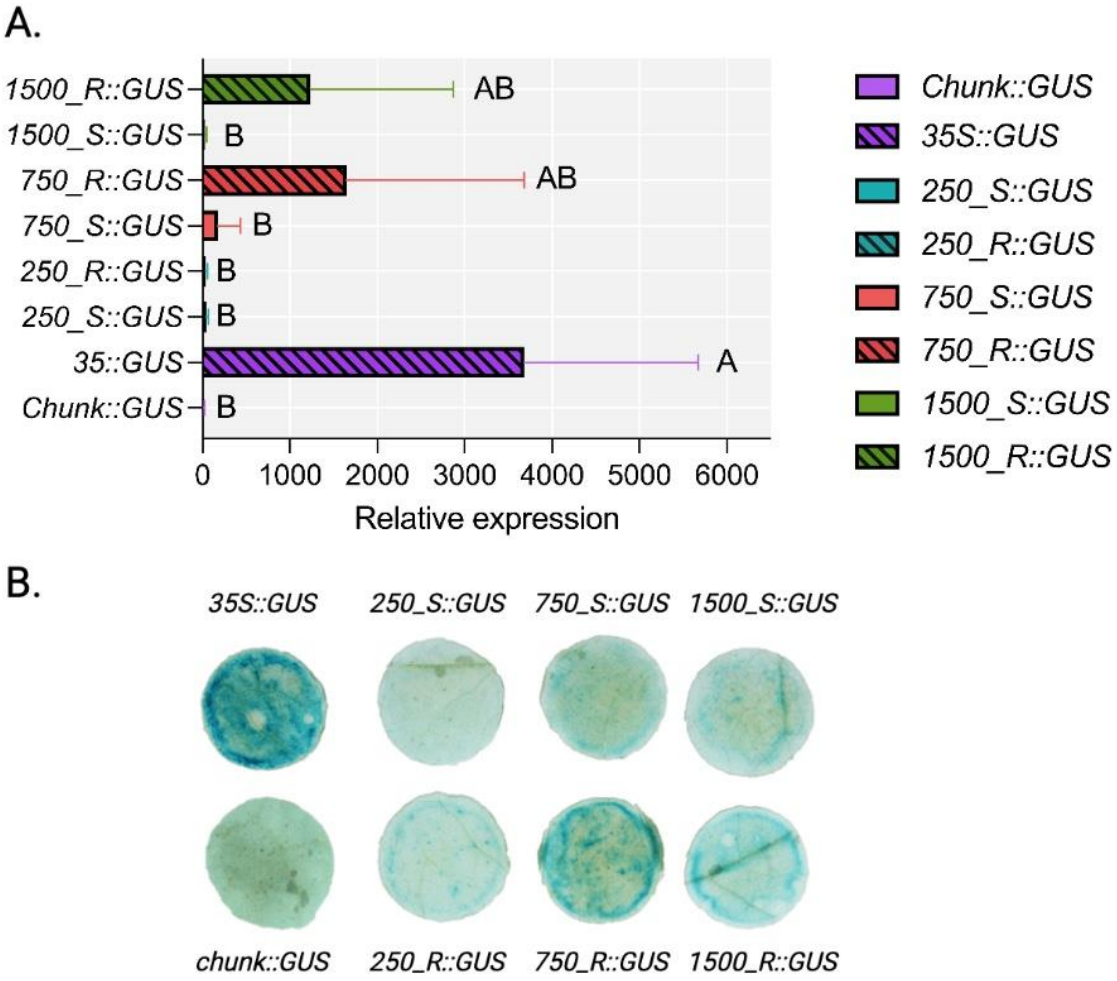


Figure 3-8. Transient expression of *GUS* and histochemical analysis in *Nicotiana benthamiana* leaf infiltrated with Agrobacterium carrying vector pBI121 with different promoters harboring the GUS gene. A) Relative *GUS* expression in *N. benthamiana* leaves infiltrated with Agrobacterium carrying the pBI121 vector with various promoters harboring the *GUS* gene. The promoters included: 35S - Cauliflower Mosaic Virus promoter; chunk - Arabidopsis chunk DNA; 250, 750, 1500 - 250, 750, 1500 bp DNA sequences of the *CYP72A219_4284* promoter from *Amaranthus palmeri* obtained from sensitive (S) or resistant (R) plants. The bar represents the standard deviation. Different letters indicate significance based on Tukey's Test ($p < 0.05$). B) Histochemical analysis of GUS expression 96 h after infiltration.

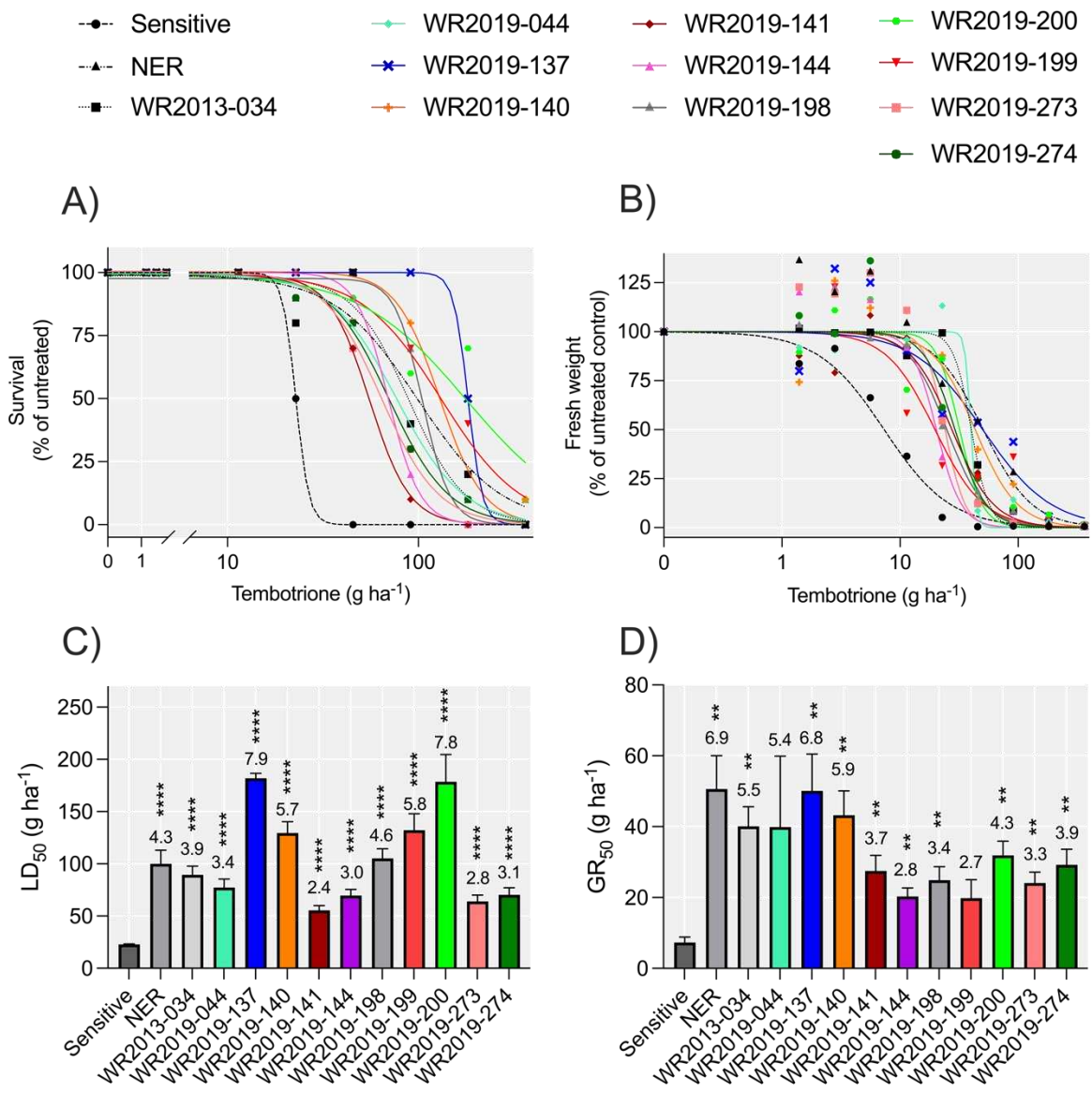


Figure 3-9. Dose response of tembotrione in different Palmer amaranth populations. A) Survival (% of untreated control) and B) Fresh shoot weight (% of untreated control) of Palmer amaranth populations submitted to increasing doses of tembotrione. Field rate 91 g a.i. ha^{-1}. B) Estimated LD_{50} and (D) GR_{50} that caused a 50 % reduction in each variable, along with the corresponding resistance index when comparing the parameter with sensitive population. Bars indicates standard error: ** <0.01, ****<0.0001.

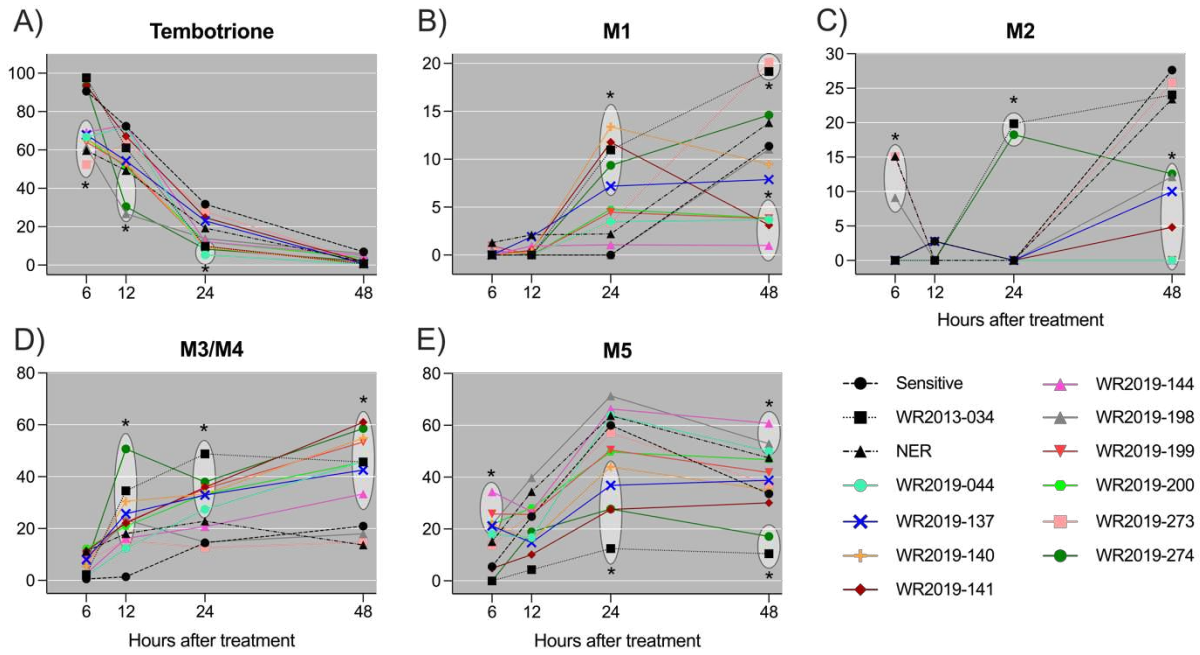


Figure 3-10. Parental tembotrione and metabolites in HPPD resistant populations of Palmer amaranth over time after tembotrione application. Average of six plants per population and treatment. A) Parental tembotrione, B) Metabolite 1, Glycosylated tembotrione; C) Metabolite 2; D) Metabolite 3 and 4 – Hydroxylated-tembotrione. E) Metabolite 5 – Reduced tembotrione. Asterisk indicates significant differences ($p<0.05$) between each population with sensitive by Dunnett's Test.

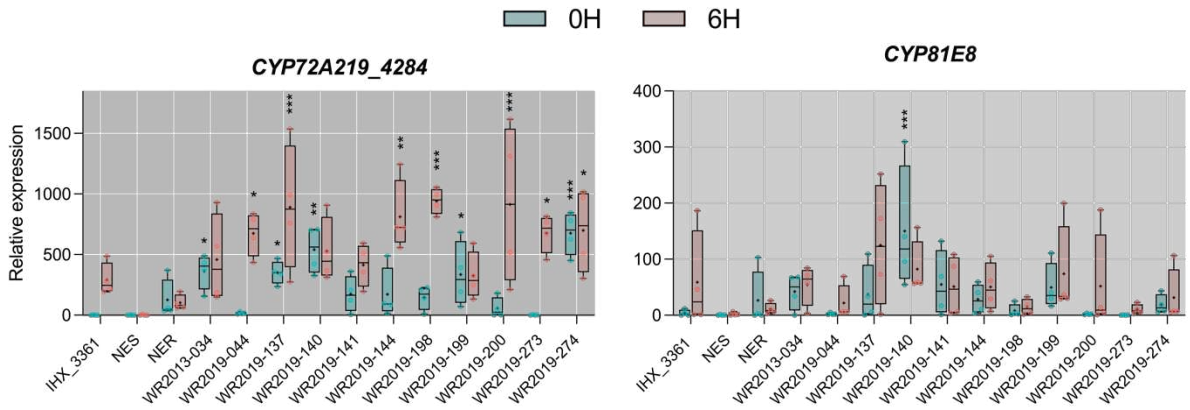


Figure 3-11. Relative gene expression of candidate cytochrome P450 genes in different Palmer amaranth population. Boxplot indicates the average, the minimum and maximum values of four biological and two technical replicates. Plants were collected before (0 h) and after (6 h) tembotrione application (91 g a.i ha^{-1}). Relative expression was calculated using the average of the normalization gene *18S* and normalized relative to the gene expression in the Sensitive biotype untreated. IHX_3361 and NES are herbicide sensitive control, NER and WR2013-034 are HPPD resistant control. The asterisk indicates the statistical difference between each population with sensitive for each time point by the Fisher's LSD test, * <0.05, ** <0.01, *** <0.001.

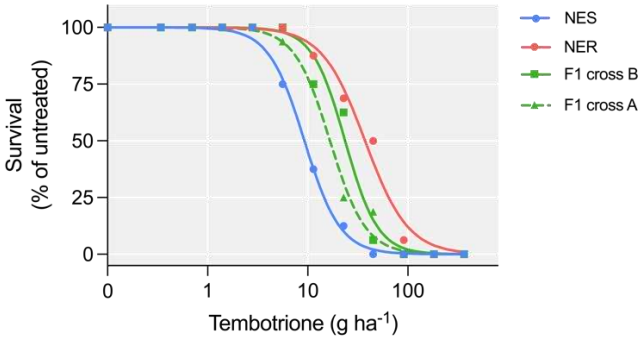


Figure 3-12. Dose-response of *Amaranthus palmeri* herbicide susceptible (NES), resistant (NER), F1 cross A and F1 cross B plants in response to tembotrione.

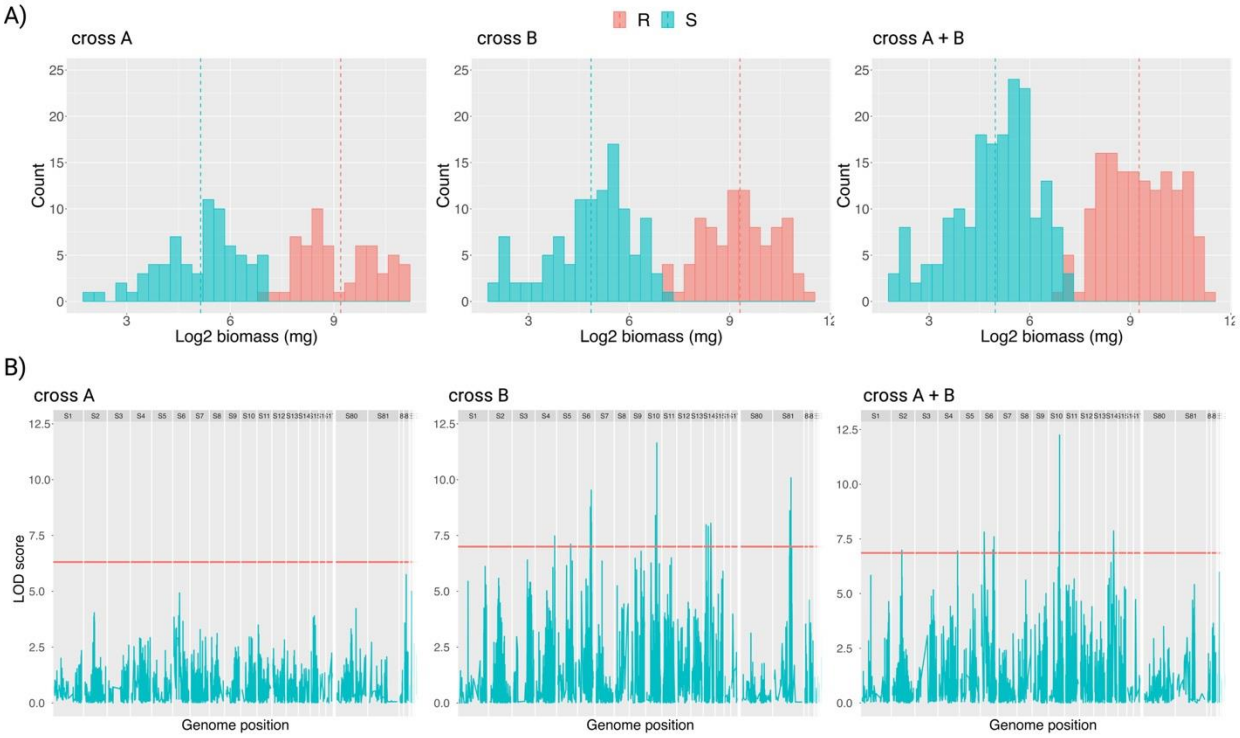


Figure 3-13. Frequency distribution and genome scan analysis. A) Distribution of fresh biomass in plants used for QTL mapping of pseudo-F2 populations from cross A, B, and combined A+B. Blue and red colors represent sensitive and resistant pseudo-F2 plants, respectively. Dashed line indicates the average biomass for parental NES (blue) and parental NER (red). B) Genome scan of pseudo-F2 Palmer amaranth populations from cross A, B, and combined A+B. The plants were treated with 77g ha^{-1} of tembotrione, and fresh biomass was measured after 28 d.

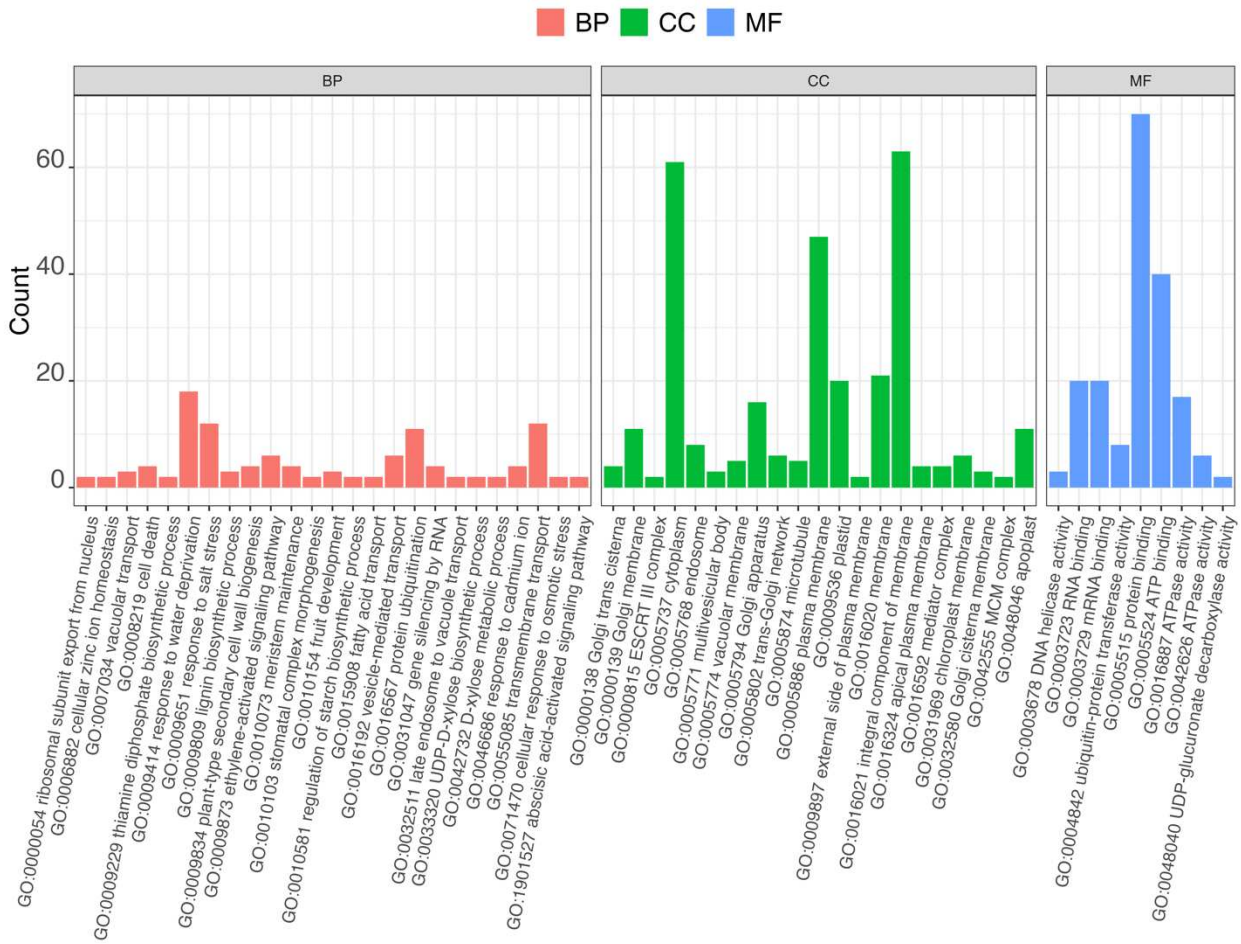


Figure 3-14. Functional enrichment of genes from QTLs of scaffold 10, scaffold 6 and scaffold 14. Detailed information of functional annotations of all the genes based on GO classification, including biological process (BP), cellular component (CC) and molecular function (MF).

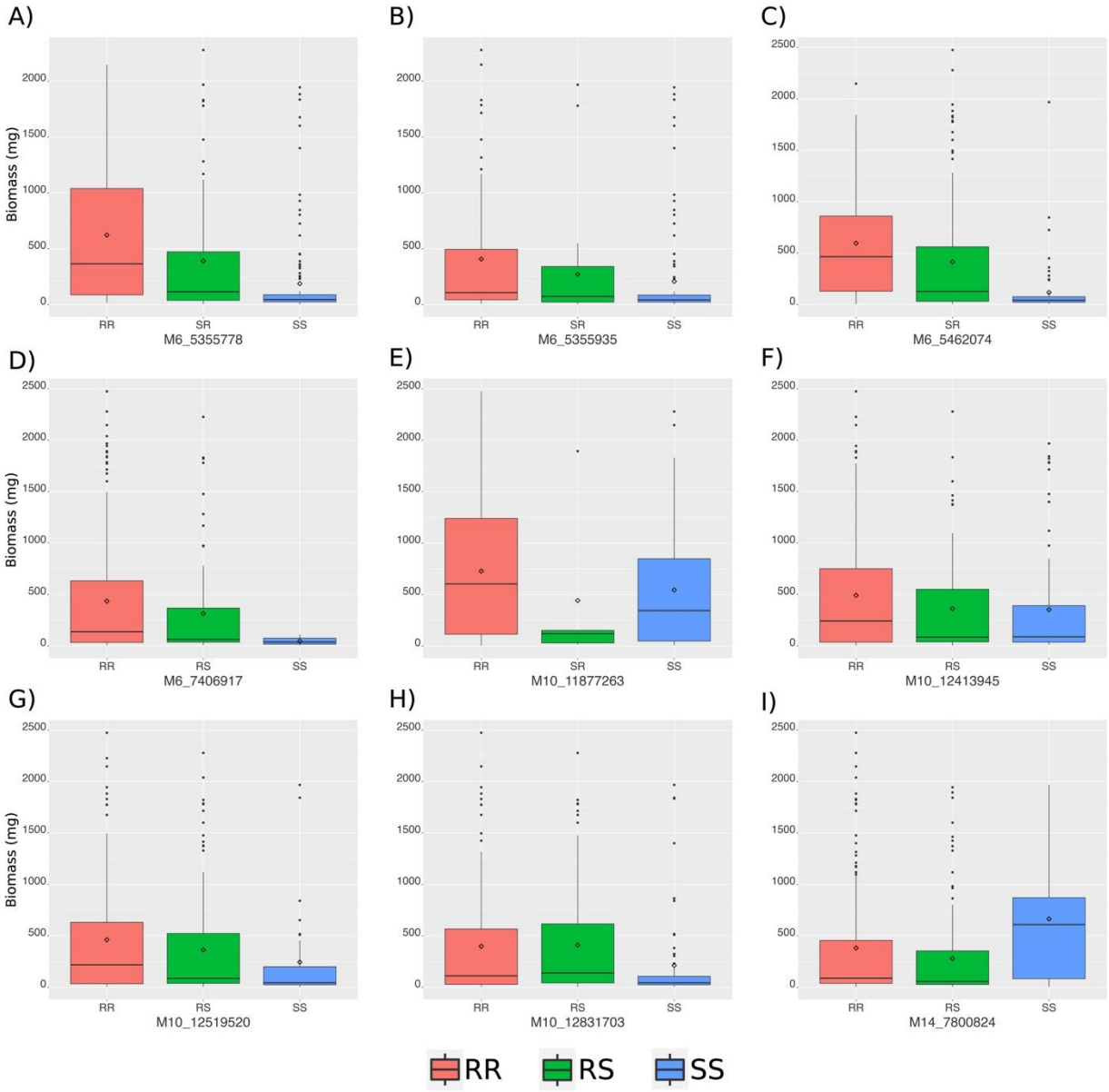


Figure 3-15. Markers associated with QTLs for tembotrione resistance in Palmer amaranth population. A - D) Markers corresponding to the QTL in scaffold 6 (M6). E - H) Markers corresponding to the QTL in scaffold 10 (M10). I) Marker found in scaffold 14 (M14). Boxplots represent the range of biomass values, including the maximum and minimum values, lower and upper quartiles, and the median. The diamond-shaped symbols indicate the average biomass for each genotype.

REFERENCES

1. Oerke EC, Crop losses to pests. *Journal of Agricultural Science* **144**: 31-43 (2006).
2. Chauhan BS, Grand challenges in weed management. *Frontiers in Agronomy* **1**: 3 (2020).
3. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A and Dayan FE, Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry* **295**: 10307–10330 (2020).
4. Rigon CAG, Gaines TA, Küpper A and Dayan FE, Metabolism-based herbicide resistance, the major threat among the non-target site resistance mechanisms. *Outlooks on Pest Manag* **31**: 162-168 (2020).
5. Kato S, Yokota Y, Suzuki R, Fujisawa Y, Sayama T, Kaga A, Anai T, Komatsu K, Oki N, Kikuchi A and Ishimoto M, Identification of a cytochrome P450 hydroxylase, *CYP81E22*, as a causative gene for the high sensitivity of soybean to herbicide bentazon. *Theoretical and Applied Genetics* **133**: 2105-2115 (2020).
6. Siminszky B, Corbin FT, Ward ER, Fleischmann TJ and Dewey RE, Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast and tobacco enhances the metabolism of phenylurea herbicides. *Proc Natl Acad Sci U S A* **96**: 1750-1755 (1999).
7. Pan G, Zhang XY, Liu KD, Zhang JW, Wu XZ, Zhu J and Tu JM, Map-based cloning of a novel rice cytochrome P450 gene *CYP81A6* that confers resistance to two different classes of herbicides. *Plant MolBiol* **61**: 933-943 (2006).
8. Pan L, Guo Q, Wang J, Shi L, Yang X, Zhou Y, Yu Q and Bai L, *CYP81A68* confers metabolic resistance to ALS and ACCase-inhibiting herbicides and its epigenetic regulation in *Echinochloa crus-galli*. *Journal of Hazardous Materials* **428**: 128225 (2022).
9. Brazier-Hicks M, Franco-Ortega S, Watson P, Rougemont B, Cohn J, Dale R, Hawkes TR, Goldberg-Cavalleri A, Onkokesung N and Edwards R, Characterization of cytochrome P450s with key roles in determining herbicide selectivity in maize. *ACS Omega* 2022).
10. Cummins I, O'Hagan D, Jablonkai I, Cole DJ, Hehn A, Werck-Reichhart D and Edwards R, Cloning, characterization and regulation of a family of phi class glutathione transferases from wheat. *Plant MolBiol* **52**: 591-603 (2003).
11. Karavangeli M, Labrou NE, Clonis YD and Tsaftaris A, Development of transgenic tobacco plants overexpressing maize glutathione *S*-transferase I for chloroacetanilide herbicides phytoremediation. *Biomol Eng* **22**: 121-128 (2005).
12. Iwakami S, Uchino A, Kataoka Y, Shibaike H, Watanabe H and Inamura T, Cytochrome P450 genes induced by bispyribac-sodium treatment in a multiple herbicide resistant biotype of *Echinochloa phyllopogon*. *Pest Manag Sci* **70**: 549-558 (2014).
13. Iwakami S, Kamidate Y, Yamaguchi T, Ishizaka M, Endo M, Suda H, Nagai K, Sunohara Y, Toki S, Uchino A, Tominaga T and Matsumoto H, *CYP81A* P450s are involved in concomitant cross-resistance to acetolactate synthase and acetyl-CoA carboxylase herbicides in *Echinochloa phyllopogon*. *New Phytologist* **221**: 2112-2122 (2019).
14. Dimaano NG, Yamaguchi T, Fukunishi K, Tominaga T and Iwakami S, Functional characterization of cytochrome P450 *CYP81A* subfamily to disclose the pattern of cross-resistance in *Echinochloa phyllopogon*. *Plant MolBiol* **102**: 403-416 (2020).
15. Zheng T, Yu X, Sun Y, Zhang Q, Zhang X, Tang M, Lin C and Shen Z, Expression of a cytochrome P450 gene from Bermuda grass *Cynodon dactylon* in soybean confers tolerance to multiple herbicides. *Plants* **11**: 949 (2022).
16. Han H, Yu Q, Beffa R, González S, Maiwald F, Wang J and Powles SB, Cytochrome P450 *CYP81A10v7* in *Lolium rigidum* confers metabolic resistance to herbicides across at least five modes of action. *The Plant Journal* **105**: 79-92 (2021).
17. Shen J, Yang Q, Hao L, Zhang L, Li X and Zheng M, The metabolism of a novel cytochrome P450 (*CYP77B34*) in tribenuron-methyl-resistant *Descurainia sophia* L. to herbicides with different mode of actions. *Int J Mol Sci* **23**: 5812 (2022).

18. Khanom S, Jang J and Lee OR, Overexpression of ginseng cytochrome P450 *CYP736A12* alters plant growth and confers phenylurea herbicide tolerance in Arabidopsis. *J Ginseng Res* **43**: 645-653 (2019).
19. Pan L, Yu Q, Han HP, Mao LF, Nyporko A, Fan LJ, Bai LY and Powles S, Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in *Echinochloa colona*. *Plant Physiol* **181**: 1519-1534 (2019).
20. Lu H, Liu Y, Li M, Han H, Zhou F, Nyporko A, Yu Q, Qiang S and Powles S, Multiple metabolic enzymes can be involved in cross-resistance to 4-hydroxyphenylpyruvate-dioxygenase-inhibiting herbicides in wild radish. *J Agric Food Chem* 2023).
21. Ward SM, Webster TM and Steckel LE, Palmer Amaranth (*Amaranthus palmeri*): A Review. *Weed Technol* **27**: 12-27 (2013).
22. Roberts J and Florentine S, A review of the biology, distribution patterns and management of the invasive species *Amaranthus palmeri* S. Watson (Palmer amaranth): Current and future management challenges. *Weed Res* **62**: 113-122 (2022).
23. Berger S, Madeira PT, Ferrell J, Gettys L, Morichetti S, Cantero JJ and Nuñez C, Palmer amaranth (*Amaranthus palmeri*) identification and documentation of ALS-resistance in Argentina. *Weed Sci* **64**: 312-320 (2016).
24. Torra J, Royo-Esnal A, Romano Y, Osuna MD, León RG and Recasens J, *Amaranthus palmeri* a new invasive weed in Spain with herbicide resistant biotypes. *Agronomy* **10**: 993 (2020).
25. Mennan H, Kaya-Altop E, Belvaux X, Brants I, Zandstra BH, Jabran K and Uysal MŞ, Investigating glyphosate resistance in *Amaranthus palmeri* biotypes from Turkey. *Phytoparasitica* **49**: 1043-1052 (2021).
26. Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE and Westra P, Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences* **107**: 1029-1034 (2010).
27. Shyam C, Peterson DE and Jugulam M, Resistance to 2,4-D in Palmer amaranth (*Amaranthus palmeri*) from Kansas is mediated by enhanced metabolism. *Weed Sci* **70**: 390-400 (2022).
28. Montgomery JS, Giacomini DA and Tranel PJ, Molecular confirmation of resistance to PPO inhibitors in *Amaranthus tuberculatus* and *Amaranthus palmeri*, and isolation of the G399A PPO2 substitution in *A. palmeri*. *Weed Technol* **35**: 99-105 (2021).
29. Jhala AJ, Sandell LD, Rana N, Kruger GR and Knezevic SZ, Confirmation and control of triazine and 4-hydroxyphenylpyruvate dioxygenase-inhibiting herbicide-resistant Palmer amaranth (*Amaranthus palmeri*) in Nebraska. *Weed Technol* **28**: 28-38 (2014).
30. Küpper A, Peter F, Zollner P, Lorentz L, Tranel PJ, Beffa R and Gaines TA, Tembotrione detoxification in 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor-resistant Palmer amaranth (*Amaranthus palmeri* S. Wats.). *Pest Manag Sci* **74**: 2325-2334 (2018).
31. Hwang J-I, Norsworthy JK, Carvalho-Moore P, Barber LT, Butts TR and McElroy JS, Exploratory analysis on herbicide metabolism and very-long-chain fatty acid production in metolachlor-resistant Palmer amaranth (*Amaranthus palmeri* S. Wats.). *J Agric Food Chem* **71**: 6014-6022 (2023).
32. González-Torralva F and Norsworthy JK, Understanding resistance mechanisms to trifluralin in an Arkansas Palmer amaranth population. *Genes (Basel)* **12** 2021).
33. Gaines TA, Lorentz L, Figge A, Herrmann J, Maiwald F, Ott M-C, Han H, Busi R, Yu Q, Powles SB and Beffa R, RNA-Seq transcriptome analysis to identify genes involved in metabolism-based diclofop resistance in *Lolium rigidum*. *The Plant Journal* **78**: 865-876 (2014).
34. Kohlhase DR, O'Rourke JA, Owen MDK and Graham MA, Using RNA-seq to characterize responses to 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor herbicide resistance in waterhemp (*Amaranthus tuberculatus*). *BMC Plant Biology* **19**: 182 (2019).
35. Murphy BP, Beffa R and Tranel PJ, Genetic architecture underlying HPPD-inhibitor resistance in a Nebraska *Amaranthus tuberculatus* population. *Pest Manag Sci* **77**: 4884-4891 (2021).

36. Chen S, Zhou Y, Chen Y and Gu J, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884-i890 (2018).
37. Montgomery JS, Giacomini D, Waithaka B, Lanz C, Murphy BP, Campe R, Lerchl J, Landes A, Gatzmann F, Janssen A, Antonise R, Patterson E, Weigel D and Tranel PJ, Draft genomes of *Amaranthus tuberculatus*, *Amaranthus hybridus*, and *Amaranthus palmeri*. *Genome Biol Evol* **12**: 1988-1993 (2020).
38. Kim D, Paggi JM, Park C, Bennett C and Salzberg SL, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**: 907-915 (2019).
39. Liao Y, Smyth GK and Shi W, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930 (2014).
40. Andersen CL, Jensen JL and Ørntoft TF, Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* **64**: 5245-5250 (2004).
41. Williams MM and Pataky JK, Genetic Basis of Sensitivity in Sweet Corn to Tembotrione. *Weed Sci* **56**: 364-370 (2008).
42. Urban P, Mignotte C, Kazmaier M, Delorme F and Pompon D, Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 *CYP73A5*. *Journal of Biological Chemistry* **272**: 19176-19186 (1997).
43. Gietz D, St Jean A, Woods RA and Schiestl RH, Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* **20**: 1425 (1992).
44. Bailey TL and Elkan C, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36 (1994).
45. Buske FA, Bodén M, Bauer DC and Bailey TL, Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics* **26**: 860-866 (2010).
46. Liu ZZ, Wang JL, Huang X, Xu WH, Liu ZM and Fang RX, The promoter of a rice glycine-rich protein gene, Osgrp-2, confers vascular-specific expression in transgenic plants. *Planta* **216**: 824-833 (2003).
47. Liu D, Shi L, Han C, Yu J, Li D and Zhang Y, Validation of reference genes for gene expression studies in virus-infected Nicotiana benthamiana using quantitative real-time PCR. *PLoS One* **7**: e46451 (2012).
48. Livak KJ and Schmittgen TD, Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**: 402-408 (2001).
49. Ritz C, Baty F, Streibig JC and Gerhard D, Dose-response analysis using R. *PLoS One* **10**: 13 (2015).
50. Bolger AM, Lohse M and Usadel B, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120 (2014).
51. Li H and Durbin R, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754-1760 (2009).
52. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAVd, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, Shakir K, Thibault J, Chandran S, Whelan C, Lek M, Gabriel S, Daly MJ, Neale B, MacArthur DG and Banks E, Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*: 201178 (2018).
53. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen Ś, Yandell BS and Churchill GA, R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics* **211**: 495-502 (2019).
54. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T and Chang W, DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* **50**: 216-221 (2022).
55. Love MI, Huber W and Anders S, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550 (2014).

56. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/. Vienna, Austria (2023).
57. Anders S and Huber W, Differential expression analysis for sequence count data. *Genome Biology* **11**: R106 (2010).
58. Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S and Werck-Reichhart D, Cytochromes p450. *Arabidopsis Book* **9**: e0144 (2011).
59. Saika H, Horita J, Taguchi-Shiobara F, Nonaka S, Nishizawa-Yokoi A, Iwakami S, Hori K, Matsumoto T, Tanaka T, Itoh T, Yano M, Kaku K, Shimizu T and Toki S, A novel rice cytochrome P450 gene, *CYP72A31*, confers tolerance to acetolactate synthase-inhibiting herbicides in rice and arabidopsis. *Plant Physiol* **166**: 1232-1240 (2014).
60. Brazier M, Cole DJ and Edwards R, O-glucosyltransferase activities toward phenolic natural products and xenobiotics in wheat and herbicide-resistant and herbicide-susceptible black-grass (*Alopecurus myosuroides*). *Phytochemistry* **59**: 149-156 (2002).
61. Iwakami S, Endo M, Saika H, Okuno J, Nakamura N, Yokoyama M, Watanabe H, Toki S, Uchino A and Inamura T, Cytochrome P450 *CYP81A12* and *CYP81A21* are associated with resistance to two acetolactate synthase inhibitors in *Echinochloa phyllopogon*. *Plant Physiol* **165**: 618-629 (2014).
62. Bourguet D and Raymond M, The molecular basis of dominance relationships: the case of some recent adaptive genes. *Journal of Evolutionary Biology* **11**: 103-122 (1998).
63. Hudson WH and Ortlund EA, The structure, function and evolution of proteins that bind DNA and RNA. *Nat Rev Mol Cell Biol* **15**: 749-760 (2014).
64. Kaundun SS, Hutchings SJ, Dale RP, Howell A, Morris JA, Kramer VC, Shivrain VK and McIndoe E, Mechanism of resistance to mesotrione in an *Amaranthus tuberculatus* population from Nebraska, USA. *PLoS One* **12**: 22 (2017).
65. Lu H, Yu Q, Han H, Owen MJ and Powles SB, Evolution of resistance to HPPD-inhibiting herbicides in a wild radish population via enhanced herbicide metabolism. *Pest Manag Sci* **76**: 1929-1937 (2019).
66. Busi R, Girotto M and Powles SB, Response to low-dose herbicide selection in self-pollinated *Avena fatua*. *Pest Manag Sci* **72**: 603-608 (2016).
67. Aarthy T, Shyam C and Jugulam M, Rapid metabolism and increased expression of *CYP81E8* gene confer high level of resistance to tembotrione in a multiple-resistant Palmer amaranth (*Amaranthus palmeri* S. Watson). *Frontiers in Agronomy* **4**2022).
68. Giacomini DA, Patterson EL, Küpper A, Beffa R, Gaines TA and Tranel PJ, Co-expression clusters and allele-specific expression in metabolism-based herbicide resistance. *Genome Biol Evol* **12**: 2267–2278 (2020).
69. Salas-Perez RA, Saski CA, Noorai RE, Srivastava SK, Lawton-Rauh AL, Nichols RL and Roma-Burgos N, RNA-seq transcriptome analysis of *Amaranthus palmeri* with differential tolerance to glufosinate herbicide. *PLoS One* **13**: e0195488 (2018).
70. Wang J, Lian L, Qi J, Fang Y, Nyporko A, Yu Q, Bai L and Pan L, Metabolic resistance to acetolactate synthase inhibitors in *Beckmannia syzigachne*: identification of *CYP81Q32* and its transcription regulation. *The Plant Journal* **n/a**2023).
71. Wang H, Xu D, Wang S, Wang A, Lei L, Jiang F, Yang B, Yuan L, Chen R, Zhang Y and Fan W, Chromosome-scale *Amaranthus tricolor* genome provides insights into the evolution of the genus *Amaranthus* and the mechanism of betalain biosynthesis. *DNA Research* **30**: 1-15 (2022).
72. Banerjee A and Roychoudhury A, WRKY proteins: signaling and regulation of expression during abiotic stress responses. *Scientific World Journal* **2015**: 807560 (2015).
73. Seale M, Hot on the Trail of DREB2A Protein Stability. *Plant Physiol* **183**: 7-8 (2020).
74. Zhang C, Hou Y, Hao Q, Chen H, Chen L, Yuan S, Shan Z, Zhang X, Yang Z, Qiu D, Zhou X and Huang W, Genome-wide survey of the soybean GATA transcription factor gene family and expression analysis under low nitrogen stress. *PLoS One* **10**: e0125174 (2015).

75. Pandolfi S and Stecca B, Luciferase reporter assays to study transcriptional activity of Hedgehog signaling in normal and cancer cells. *Methods Mol Biol* **1322**: 71-79 (2015).
76. Reece-Hoyes JS and Marian Walhout AJ, Yeast one-hybrid assays: a historical and technical perspective. *Methods* **57**: 441-447 (2012).

4. CHAPTER III - UNRAVELING THE ROLE OF P450 REDUCTASE IN HERBICIDE METABOLIC RESISTANCE MECHANISM

INTRODUCTION

Metabolic resistance is one of the most important mechanisms of herbicide resistance in weeds. Herbicide metabolism occurs in weeds that were selected by herbicides with enhanced capacity to metabolize herbicides before reaching the target enzyme[1]. This mechanism of resistance threatens weed management due to its broad substrate specificity to many herbicides from different mechanisms of action[2]. Metabolic herbicide resistance is divided in three phases, known as Phase I – oxidation, Phase II – conjugation, and Phase III – sequestration and degradation. Phase I is mediated primarily by cytochrome P450 monooxygenase (P450) enzymes. These enzymes are membrane-bound, localized in the endoplasmic reticulum, and are involved in the biosynthesis of many molecules including antioxidants, secondary metabolites, and in detoxification of many xenobiotic compounds[3]. P450s are the most important in phase I herbicide metabolism. Plant genomes have a high number of P450 genes that play a crucial role in herbicide resistance in metabolic resistant weed species[2].

Cytochrome P450 reductase (CPR) is a membrane-bound enzyme localized in the endoplasmic reticulum that plays a critical role in the biosynthesis of a wide range of compounds in living organisms, including plants, animals, and bacteria. The structure of P450 reductase is divided in four domains, including N-terminal FMN-binding domain, a connecting domain, an FAD-binding domain, and a C-terminal NADPH-binding domain[4, 5] (Figure 4-1). CPR enzymes arose from of a fusion of different genes encoding FAD-binding, ferredoxin-$NADP^+$, and an FMN-flavodoxin[6]. Living organisms require CPR to supply two electrons for P450 enzymes to conduct each monooxygenase reaction with the organic substrate[7, 8]. The CPR reduction includes

multiple steps for electron transfer. The first step is reduction from NADPH to FAD, then FMN, and finally to the P450 acceptor[9] (Figure 4-1). The transfer of electrons to the P450 active site in the P450 catalytic cycle relies on a robust and intricate CPR:P450 complex in the membrane[10]. This function of CPR makes it an essential component of specialized metabolism.

In contrast to the huge number of P450 genes within plant genomes, for example, 246 in *Arabidopsis thaliana* (Arabidopsis)[11], 328 in *Oryza* spp. (rice)[12], 867 in *Echinochloa crus-galli* (barnyardgrass) [13], 694 in *E. colona* (jungle rice)[14], and 323 in *Conyza canadensis* (horseweed) [15], CPRs generally have only two or three copies in plant genomes. There are two CPR genes in Arabidopsis, referred to as *Arabidopsis thaliana* cytochrome P450 reductase 1 (*ATR1*), 2 (*ATR2*) [16, 17] and one hypothetical reductase 3 (*ATR3)* that encodes a diflavin reductase essential for embryogenesis, but that is unable to reduce P450 *in vitro*[18]. The isoforms are grouped in two distinct groups with different roles. Class I is *ATR1*, which is involved in primary and basal metabolism. Class II is *ATR2*, which is inducible by environmental stimuli such as wounding, pathogen attack, or high light intensity exposure[19, 20].

The hypothesis of this study was that *ATR1* and *ATR2* are the main source of electrons for P450 enzymes and that silencing these genes would decrease the activity of P450s that are involved in herbicide metabolism. The main objective was to evaluate the function of *ATR1* and *ATR2* in transgenic Arabidopsis over-expressing the P450 *CYP81A12*.

MATERIALS AND METHODS

Plant material

Mutant SALK lines of Arabidopsis were ordered from the Arabidopsis Biological Resource Center (ABRC - https://abrc.osu.edu/). The lines SALK_111558, SALK_208483C have a T-DNA insertion in Arabidopsis cytochrome P450 reductase 1 (*ATR1* - at4g24520), hereafter named as *atr1-a* and *atr1-b*, respectively. The lines SALK_026053 and SALK_152766 have T-DNA insertion on Arabidopsis cytochrome P450 reductase 2 (*ATR2* - at4g30210), hereafter named as *atr2-a* and *atr2-b*.

Seeds from mutants and wild-type Columbia_0 (Col_0) were sown in soil and kept in a growth chamber at 22°C/18°C with a light intensity of 120 μE m^{-2} s^{-1} for 16 h. When seedlings had four true leaves, they were transplanted into individual pots and kept in the growth chamber in the same conditions. All the following experiments used the same methods to grow the Arabidopsis plants.

Characterization and confirmation of T-DNA insertion

PCR assays were performed to identify and confirm the presence of T-DNA insertion in the *ATR1* and *ATR2* genes in the mutant lines and to verify the homozygosity. To amplify the wild type (WT) version of *ATR1* or *ATR2* genes, different primers set named as LP (forward) and RP (reverse) spanning the T-DNA insertion specific for each gene and line were used. The primers to identify the T-DNA insertion for all the mutant lines were RP primers, specific for each gene and line, with LBb1.3 (Figure 4-2). The primer LBb1.3 is a primer designed in the right border of the T-DNA sequence (Figure 4-2). Primer sequences of LBb1.3, LP and RP for *atr1-a*, *atr1-b* and *atr2-a* were provided by SIGnAL (http://signal.salk.edu/). Primer set for *atr2-*

b were designed using Primer3plus. The primers sequences and their characteristics are described in Table 4-1.

Leaf tissue was collected from individual plants of mutants and Col_0 and incubated in 1.2 mL tubes when the plants were at rosette stage. Around 20 plants for each mutant line were genotyped. The tissues were frozen at -20°C until further analyses. Frozen materials were ground using a Qiagen Tissuelyzer and DNA was extracted using a modified hexadecyltrimethylammonium bromide (CTAB) method. The quality and quantification of DNA were assessed using Nanodrop2000.

PCR reactions were performed with Econotaq Plus 2× Master Mix. PCR reaction consisted of 12.5 µL of Econotaq Master Mix, 6.5 µL nuclease-free water, and 2.5 µL of forward and reverse primers (10 µM). Each sample was submitted to two different reactions, called "WT PCR" and "T-DNA PCR". The former consisted of LP and RP primers to amplify the wild-type form or heterozygous and the latter was composed by RP and LBb1.3 primers to amplify the T-DNA insertion if present in the sample. PCR cycling conditions were composed by initial denaturation of 94°C for 2 min, followed by 35 cycles consisting of denaturation of at 94°C for 30s, annealing at 55°C for "WT PCR" and 48°C for "T-DNA PCR" for 30 s, followed by extension at 72°C for 1:10 min, with a final extension time at 72°C for 5 min. Schemes of the primers and PCRs are represented in Figure 4-2. The DNA amplification was confirmed by running in 1% agarose gel electrophoresis and PCR products were sent for Sanger sequencing to confirm the T-DNA insertion.

Generating transgenic herbicide-resistant Arabidopsis with mutant *atr1* or *atr2* genes

Aiming to generate a transgenic Arabidopsis line with a loss of function in *ATR1* or *ATR2*, a transgenic Arabidopsis line carrying the P450 *CYP81A12* from *Echinochloa phyllopogon* (or *E.*

oryzicola, late watergrass) was provided by Dr. Iwakami from Kyoto University. The transgenic Arabidopsis metabolizes five different herbicide modes of action including acetolactate synthase (ALS) inhibitors (bensulfuron-methyl, pyrazosulfuron-ethyl, chlorsulfuron, azimsulfuron, imazamox, propoxycarbazone-sodium and penoxsulam), 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor (mesotrione), protoporphyrinogen oxidase (PPO) inhibitor (pyraclonil), deoxyxylulose phosphate synthase (DXS) inhibitor (clomazone), and photosystem II (PSII) inhibitor (bentazon)[21]. This population hereafter is called Arabidopsis *CYP81A12*.

Transgenic Arabidopsis *CYP81A12* and mutant lines *atr1-a*, *atr1-b*, *atr2-a* and *atr2-b* were grown with the same conditions as indicated before. When plants were in the flowering stage, around four weeks after transplanting, the mutant lines were crossed with the transgenic Arabidopsis *CYP81A12*. The crossing consisted of cutting out the stamens from the immature flowers of Arabidopsis *CYP81A12* and in the next day pollinating the carpel with pollen from different mutant lines flowers. The cross was repeated more three times with the new flowers within an interval of a week. The crossed flowers were labeled, and the seeds were collected when siliques turned brown. In total, four crosses were performed, cross A – transgenic *CYP81A12* × *atr1-a*, cross B – transgenic *CYP81A12* ×*atr1-b*, cross C – transgenic *CYP81A12* ×*atr2-a*, and cross D – transgenic *CYP81A12* × *atr2-b* (Figure 4-3).

F1 plants were grown, and tissue collected at 2 weeks after transplanting for DNA extraction. Extracted DNA was used to genotype for T-DNA insertion and presence of *CYP81A12* gene. T-DNA insertion was genotyped as indicated before, using two PCR reactions, WT PCR and T-DNA PCR. PCR amplification for the whole *CYP81A12* was performed to genotype for the presence of the gene using genomic DNA. Primer sequences for *CYP81A12* are listed in Table 4-1. Plants that were heterozygous for T-DNA insertion and had the presence of

CYP81A12 were selected and self-pollinated. Seeds were collected when mature. F2 plants were grown with the same conditions as described before, and tissue collected for genotyping. Genomic DNA was extracted, and T-DNA insertion was genotyped using the same two PCR reactions as indicated before. It was expected that 25% of the F2 population were homozygous for T-DNA insertion. These plants were selected for genotyping the presence of *CYP81A12* by gene amplification and those plants that have the gene were selected. These plants were genotyped for copy number analysis of *CYP81A12* using QX200 Droplet Digital PCR system (ddPCR) due to its precision and sensitivity for detection of low copy numbers. ddPCR consisted of a master mix composed of 10 µL of 2×QX200 EvaGreen® Supermix, 2.5 µL of each primer (5 µM), and 5 µL of genomic DNA (5 ng/µL).

Selected samples were processed for droplet generation using QX200™ droplet generator, followed by PCR reaction with initial cycle of 5 min at 95°C followed by a sequence of 40 cycles starting at 95°C for 30 s and 60°C for 1 min. Droplet reading of PCR products was conducted in the QX200 Droplet Reader. The *ALS* gene was used as a reference gene containing only two copies (one copy per haplotype). Primer sequences for *ALS* and *CYP81A12* used for ddPCR are listed in Table 4-2. The samples were analyzed by QuantaSoft AP Software, which measures the number of positive and negative droplets for each gene in each well. Droplets are assigned as positive or negative by thresholding based on their fluorescence amplitude. The software calculates the starting concentration of each target DNA molecule by modeling a Poisson distribution; the formula used for Poisson modeling is: Copies per droplet = -ln(1-p) where p= fraction of positive droplets. Concentrations are provided in units of copies per microliter of input sample with 95% confidence intervals. The ratio of the number of copies per droplet of *ALS* and *CYP81A12* was calculated. After genotyping, the Arabidopsis transgenic

CYP81A12 carrying mutant *atr1* or *atr2* hereafter are named as *CYP81A12 atr1-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a* and *CYP81A12 atr2-b*.

Generating double knockout *atr1 atr2* Arabidopsis lines

In order to obtain an Arabidopsis line with a double knockout of *ATR1* and *ATR2* within the same plant, mutant lines *atr1-a*, *atr1-b*, *atr2-a*, and *atr2-b* were cultivated under the same specified conditions as previously indicated. Upon reaching the flowering stage, crosses were performed as described earlier. F1 plants were subjected to genotyping to determine homozygosity for *ATR1* and *ATR2* using the same PCR reactions as previously mentioned. Heterozygous plants were selected for self-pollination. F2 seeds were then sown in soil, and subsequent plants were genotyped for *atr1* and *atr2* mutations.

Gene expression analysis in *atr1* and *atr2* mutant lines

Mutant and transgenic mutant lines were grown in a growth chamber with the same conditions as specified before. Approximately 100 mg of leaf tissue were collected from plants in the rosette stage, around two weeks after transplanting. Sampling consisted of cutting the youngest fully expanded leaf for each biological replication, placing in 2 mL tubes, and incubating immediately in liquid nitrogen and afterwards kept at -80°C. Plants of lines Col_0, transgenic *CYP81A12*, and transgenic mutants *CYP81A12 atr1-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a* and *CYP81A12 atr2-b* were kept for fresh mass measurement at 28 d after transplant to analysis any fitness cost in the mutant lines.

RNA extraction was conducted using Direct-zol™ RNA miniprep from Zymo Research following manufacturer instructions. DNase I treatment was conducted during the RNA extraction protocol. Quantification and quality control were performed using Nanodrop 2000. cDNA synthesis was conducted using iScript™ from Zymo Research using 1 µg of RNA. Gene

expression was analyzed on CFX96 Real Time System using SsoAdvanced™ Universal SYBR® Green Supermix. The reactions consisted of 5 µL of cDNA from samples diluted in Nuclease-Free Water in 1:20, 2.5 µL of each primer at 5 µM and 5 µL of SsoAdvanced universal SYBR® Green supermix (2×). Thermal cycler conditions consisted of an initial cycle of 30 s at 95°C, followed by a sequence of 35 cycles starting at 95°C for 10 s and 72°C for 30 s. Melt-curve analysis was performed by adding steps incrementing temperature 65–95°C by 0.5°C.

Primer sets were specifically designed to target various regions of the genes, both upstream and downstream of the T-DNA insertion site. This systematic design allowed for accurate analysis of how the gene was affected by the T-DNA insertion, facilitating the identification of gene knockout. For *ATR1* gene, four sets of primers were designed, called ATR1_A with forward primers in the 5'UTR and reverse primers in the 1st exon (spanning the T-DNA insertion position in *atr1-a* line), ATR1_B with forward and reverse primers in the 1st exon, ATR1_C with the forward primer in exon 4th and reverse primer in exon 5th (spanning the T-DNA insertion position in *atr1-b* line) and ATR1_D with forward in the 8th exon and reverse in the 9th exon (Figure 4-4). For *ATR2* gene, five sets of primers were designed, ATR2_A with forward and reverse primers localized in the 1st exon, ATR2_B with forward primer in the 3th exon and reverse primer in the 4th exon (spanning the T-DNA insertion position in *atr2-a* line), ATR2_C with forward primer localized in the 9th and reverse in the 10th exon, ATR2_D with forward primer in the 12th exon and reverse in the 13th (spanning the T-DNA insertion position in *atr2-b* line) and ATR2_E with forward and reverse primers designed in the 14th exon. The primer location in *ATR1* and *ATR2* are represented in Figure 4-4 A and B, respectively. Primer sequences for gene expression and their characteristics are listed in Table 4-2.

Herbicide sensitivity of Arabidopsis lines in ½ MS medium

Plant growth of different lines of Arabidopsis were evaluated on ½ Murashige and Skoog (MS) solid medium containing different doses of bensulfuron, chlorsulfuron, penoxsulam, propoxycarbazone, mesotrione and bentazon. Seeds of Arabidopsis wild type (Col_0), transgenic Arabidopsis *CYP81A12*, mutant *atr1-a*, *atr1-b*, *atr2-a*, *atr2-b* and transgenic mutant *CYP81A12 atr1-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a*, *CYP81A12 atr2-b* were sterilized by chloride gas for 7 h by mixing 100 mL of bleach with 3% hydrochloric acid. Technical grade herbicides were diluted in ethanol (bensulfuron, penoxsulam, propoxycarbazone, bentazon) or methanol (chlorsulfuron, mesotrione) to mix with ½ MS medium before it solidified. Petri dishes containing ½ MS with different herbicides doses were divided by 10 sections, where six seeds of each Arabidopsis line were placed. The plates were incubated in growth chamber at 25°C, 14-h photoperiod and 110 μE m^{-2} s^{-1} light intensity until the day of evaluation. The number of green leaves were counted 14 d after plating, and the data were normalized to percentage of untreated control.

Whole-plant dose response

Arabidopsis plants of Col_0, transgenic *CYP81A12* and transgenic mutant *CYP81A12 atr1-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a*, *CYP81A12 atr2-b* were grown with same conditions as specified before. Plants in rosette stage (14 d after transplanting) were sprayed with increasing doses of Chlorsulfuron (Telar ® XP, 750 g Kg^{-1}), penoxsulam (Ricer, 217 g a.i. kg^{-1}), mesotrione (Callisto 480g a.i. L^{-1}) and 2,4-D (Shredder 2,4-D LV4, 455g a.e. L^{-1}). The experimental design was completely randomized in 6 × 11 with four replications. Factor A was Arabidopsis lines as described above. Factor B was increasing herbicides dose for each herbicide. Chlorsulfuron and penoxsulam doses were 0, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, 12.8, 25.6 and 51.2 g a.i.; mesotrione doses were 0, 3, 6, 12, 25, 50, 100, 200 and 400 g a.i.; and 2,4-D

doses were 0, 1, 2, 4, 8, 16, 32, 64, and 128 g a.e. The shoot fresh mass was measured at 28 d after herbicide treatments by cutting the plant close to the soil surface.

Herbicide metabolism analysis

Seedlings of Arabidopsis lines Col_0, transgenic *CYP81A12*, and transgenic mutant *CYP81A12 atr-a*, *CYP81A12 atr-b*, *CYP81A12 atr2-a* and *CYP81A12 atr2-b* were transplanted into pots filled with foil and grown under same conditions as cited before. At rosette stage, mesotrione (Calisto ®, 480 g L^{-1} mesotrione, Syngenta, Switzerland) or chlorsulfuron (Telar ® XP, 750 g Kg^{-1} chlorsulfuron, Bayer CropScience, Germany) were applied in an automated spray chamber (Greenhouse Spray Chamber, model Generation IV) using a TJ8002E nozzle, calibrated to deliver 200 L ha^{-1} at a pressure of 280 kPa and speed of 1.2 m s^{-1}. Herbicides were applied at 100 g ha^{-1} for mesotrione and 50 g ha^{-1} for chlorsulfuron. The doses were chosen based on the whole-plant dose-response curve experiment described above as a good dose to discriminate the resistant and sensitive plants. After herbicide application the plants were kept in the growth chamber at same conditions and whole plants were collected at 6, 12, 24, 48, and 96 h for mesotrione and adding 144 h and 196 h for plants treated with chlorsulfuron. At each time of tissue collection, plant weight was measured and washed with 80% acetone for three times, dried out with towel paper and incubate in 15 mL Falcon tubes and kept at –70°C ultra-freezer until the extraction. The extraction was conducted by grinding the tissue with a pestle, then adding ten times of the plant's weight with 80% methanol and 0.1% formic acid. The solution was homogenized for 30 s using PowerGen 125 (Fisher Scientific) and centrifuged for 10 min at 4,700×g (Sorvall Legend X1R (Thermo Scientific, Waltham, MA). The supernatant was filter with Econofltr Nylon 13 mm 0.2 µm (Agilent Technologies, Santa Clara, CA) and injected in the LC/MSMS.

LC-MS/MS system consisted of a Nexera X2 UPLC with 2 LC-30AD pumps, a SIL-30AC MP autosampler, a DGU-20A5 Prominence degasser, a CTO-30A column oven, and SPD-M30A diode array detector coupled to an 8040 quadrupole mass-spectrometer (Shimadzu Scientific Instruments, Columbia, MD). For mesotrione, the MS was in positive mode [M+H]+ with MRM optimized for mesotrione 340>227.95 and set for 100 ms dwell time with a Q1 pre-bias of -16.0V, a collision energy of -18.0V and a Q3 pre-bias of -16.0V and for hydroxy-mesotrione 356>55>228 and set for 100 ms dwell time with a Q1 pre-bias of -16V, a collision energy of -18.0V and a Q3 pre-bias of -16.0V. For chlorsulfuron, the MS was in positive mode with MRM optimized for chlorsulfuron 357.9>141.1>167.05 and set for 100 ms dwell time with a Q1 pre-bias of -25.0V, a collision energy of -20.0V and a Q3 pre-bias of -15.0V and for hydroxy-chlorsulfuron 374.0>127.0>182.0 and set for a collision energy of -35.0. The samples were chromatographed on a 100×4.6 mm kinetex 2.6 µm biphenyl column (Phenomenex, Torrance, CA) maintained at 40°C. For both herbicides, solvent A consisted of water with 0.1% formic acid and solvent B was methanol with 0.1% formic acid. The solvent program started at 50% B and increased to 95% B in 5 min and maintained at 95% for 4 min. The solvent was returned to 50% B and maintained there for 3 min before the next injection. The flow rate was set at 0.4 mL min^{-1} and each sample were analyzed with 5 µL injection volumes.

Identification and phylogenetic analysis of CPRs in different weeds

The presence of *CPRs* and their copies was searched in several important weed species. The genomes of *Alopecurus myosuroides*, *Chenopodium formosanum*, *Bassia scoparia*, *Eleusine indica*, *Bromus tectorum*, *Echinochloa crus-galli*, *E. oryzicola*, *E. colona*, *Ipomoea purpurea, Poa annua* and *Conyza canadensis* were accessible through WeedPedia (https://weedpedia.weedgenomics.org). CPRs were identified with a blast search using *ATR1*

(AT4G24520) or *ATR2* (AT4G30210) against the genomes or utilized the InterPro ID for NADPH-cytochrome P450 reductase (IPR023208). Additionally, the genome of *Amaranthus palmeri* (Palmer amaranth) and *A. tuberculatus* (common waterhemp) were available, and the COGE web browser (https://genomevolution.org/coge) was used for blasting purposes. The *CPRs* in *Amaranthus* species were identified using a comparative analysis approach. Firstly, *ATR1* and *ATR2* sequences were subjected to a sequence similarity search against the *Amaranthus hypochondriacus* (Prince-of-Wales feather) genome, which is accessible through the Phytozome database. Subsequently, the identified sequences were further compared against the Palmer amaranth and common waterhemp genomes. Other already known plant *CPRs* were extracted from genbank NCBI, such as corn (*Zea mays*, CPR2a: NP_001159331.1, CPR2b: NP_001146345.2), rice (*Oryza sativa*, CPR2a: XP_015651232.1, CPR2b: XP_015650780.1), soybean (*Glycine max*, CPR1: NP_001236742.2, CPR2: XP_003549436.1), cotton (*Gossypium hirsutum*, CPR1: ACN54323.1, CPR2: NP_001314398.2), and Barrel medic (*Medicago truncatula*, CPR1: XP_003602898.1, CPR2: XP_003610109.1). Multiple protein alignment was performed using ClustalO and used for tree construction with Neighbor-joining method using Geneious Prime® 2023.0.1. The N-terminal, transmembrane, FMN binding, FAD binding, and NADP binding domains were identified. The transmembrane was identified using DeepTMHMM tool[22] (https://dtu.biolib.com/DeepTMHMM).

Statistical analysis

Gene expression analysis of *ATR1* and *ATR2* gene in mutant plants was performed using *ALS* as reference gene due to its stability with different conditions and herbicide applications[23]. Expression analyses were performed utilizing four biological and two technical replicates for each line. The mean Ct values and the standard deviation were calculated by treatment. A melt

curve analysis confirmed the presence of a single amplified product for each reaction, based on presence of a single melting temperature consistent across samples. Relative transcript abundance was calculated using $2^{(-\Delta\Delta Ct)}$ method for each gene and biotype[24]. The reference population used for the calculation was Col_0. Dunnett's test ($p < 0.05$) was used to compare the relative expression between Col_0 and mutant lines.

For the dose-response experiments *in vitro* and *in planta*, the response variables evaluated were number of green leaves at 14 d after seeds were planted and shoot fresh weight (% of untreated control) at 28 d after herbicide application (DAA) for the whole-plant dose-response experiment. Data were converted into percentages relative to untreated treatment for both experiments. The drc package [25] from statistical software R v.3.5.3[26] was used for data analysis. The data were tested for the best nonlinear model by comparing different models by *ANOVA()* and using function *modelFit ()* testing for lack-of-fit. Based on the analysis, the data were adjusted using the three-parameter log-logistic model with the function *modelFit ()* from the drc package[25]. For the whole-plant dose response experiment, the herbicide doses causing 50% reduction in shoot fresh weight (GR_{50}) were determined using the function *summary()* and the resistance index (RI) was calculated by the function *EDcomp()* by dividing the GR_{50} from transgenic *CYP81A12* or transgenic *CYP81A12* mutants (*atr1* or *atr2*) by the GR_{50} of Col_0. For the dose-response in plates the calculation was the same, using the GR_{50} obtained by the variable green leaves analyzed. The x and y data from the dose-response curves of both experiments were extracted from R using the function *write.csv()* and plotted with the treatment means by GraphPad Prism version 9 for Windows (San Diego, California).

The metabolism data obtained by LC/MSMS were submitted to analysis of variance (ANOVA) and the means were compared using Dunnett's test ($p < 0.05$). The peak area of the herbicides and metabolites of all the different Arabidopsis lines were compared to the Col_0 line.

RESULTS

Arabidopsis mutants

SALK lines *atr1-a*, *atr1-b* and *atr2-a* were segregating for T-DNA insertion and heterozygous plants were selected to self and seed kept for seed production. *atr2-b* was homozygous for T-DNA insertion and four plants were selfed and produce seeds. Progeny plants from *atr1-a*, *atr1-b* and *atr2-a* were genotyped again and individuals homozygous for the T-DNA insertion were selected to self and seeds were produced. Homozygous T-DNA insertion in *ATR1* in *atr1-a* and *atr1-b* lines and homozygous T-DNA insertion in *ATR2* in *atr2-a* and *atr2-b* lines were confirmed by PCR. Primers LP and RP for both genes amplified only when Col_0 DNA was in the reaction. No SALK line amplified the WT version of the respective gene and only amplified the T-DNA reaction, confirming the homozygosity for T-DNA insertion. The amplifications were confirmed by electrophoresis in 1% agarose gel (Figure 4-5), followed by Sanger sequencing. Based on Sanger sequencing results, the insertion of T-DNA was in the 5' UTR and in the 4th intron in *atr1-a and atr1-b* mutant line, respectively, for *ATR1* gene, and in the 4th intron and the 12th exon in *atr2-a* and *atr2-b* mutant line, respectively, for *ATR2* gene (Figure 4-5).

Transgenic Arabidopsis *CYP81A12* with *atr1* and *atr2* knockdown

F1 plants from crosses were genotyped for T-DNA insertion. Around 20 plants were genotyped from each cross and around 90% of the plants were heterozygous for *ATR1* T-DNA

insertion for cross A and B and heterozygous for *ATR2* T-DNA insertion in cross C and D, indicating success in the crosses. These plants were genotyped for *CYP81A12* presence by PCR and plants that had the gene were selected for self-pollination to generate F2. F2 plants segregated 1:2:1 for *ATR1* T-DNA insertion for cross A and B and for *ATR2* T-DNA insertion for cross C and D (Table 4-3). Homozygous mutants were submitted for PCR amplification of *CYP81A12* and when confirmed, samples were submitted for ddPCR. Four, one, three, and five plants out of 15, 12, 9 and 15 plants homozygous had *CYP81A12*. These plants were analyzed by ddPCR. When only one gene copy was identified in comparison to *ALS* gene, the plants were self-pollinated, and their progenies were analyzed again using ddPCR.

The ddPCR results indicated the same copy number of *CYP81A12* and *ALS* for the parental transgenic line *CYP81A12* (Table 4-4). Plants selected based on previous genotyping (as indicated above from crosses A, B, C, and D) had a ratio of *ALS* and *CYP81A12* from 0.9 to 1.1, suggesting two copies of each gene, one copy per haploid genome. These plants were selected and used for further experiments.

atr1 and *atr2* double knockouts

Genotyping analysis of F2 plants obtained from various crosses between *atr1* and *atr2* mutants demonstrated an absence of double knockout genotypes (Supporting information Table 4-1). A total of 416 F2 plants were genotyped, revealing seven out of the nine possible genotypes. Among the genotyped plants, 46.2% exhibited heterozygosity for both genes, 5% displayed heterozygosity for one gene and wildtype for the other, 6.2% exhibited homozygous mutation in one gene and heterozygosity in the other, while 42.6% exhibited homozygous mutation in one gene and wildtype in the other. Notably, double wildtype and double mutant genotypes were not observed. The segregation pattern observed in the F2 generation deviated

from the expected Mendelian segregation for two independent genes, indicating the presence of linked genes. Further investigation revealed that the *atr1* and *atr2* genes are localized on chromosome four of Arabidopsis and are approximately two megabases (MBp) apart, providing supporting evidence that the two genes may be linked.

ATR1 and *ATR2* gene expression in mutant and transgenic mutant lines

Different primer sets were able to detect the loss of function of *ATR1* and *ATR2* genes in the mutant lines. Mutant *atr1-a* showed a reduced transcription of *ATR1* when the set primer used was spanning the T-DNA insertion (Figure 4-6A); however, normal expression of *ATR1* gene was observed when primer set used was downstream of the T-DNA insertion (Figure 4-6B-D), indicating that the T-DNA insertion in 5' UTR did not decrease *ATR1* expression in *atr1-a* line. In other hand, mutant line *atr1-b*, which has T-DNA insertion in the 4th intron was able to reduce significantly the *ATR1* transcripts. It was confirmed by disrupting the *ATR1* transcript when the primer set used was spanning the T-DNA insertion and downstream it (Figure 4-6B-D). Interesting that this line had an increased transcript when set of primers used was upstream the T-DNA (Figure 4-6A). These results indicate some tentative evidence of compensation of the gene expression of the disrupted gene by the plant.

Mutants for *ATR2* showed similar results. Mutant *atr2-a* had an upregulation of transcript upstream of the T-DNA insertion in this line (Figure 4-6E); however, this line had a significant reduction of *ATR2* expression downstream of the T-DNA insertion, especially when the set of primers used was spanning the T-DNA insertion (Figure 4-6F-I). Mutant *atr2-b* has T-DNA insertion in the 12th exon and had a reduction in the transcription upstream and downstream of the T-DNA insertion (Figure 4-6E-I). T-DNA insertion in the *ATR1* mutant *atr1-b* and in the *ATR2* mutants *atr2-a* and *atr2-b* had knockdown of expression of their respective mRNA.

CYP81A12 expression was analyzed, and all the transgenic mutant line had the same transcription production as the transgenic *CYP81A12* (Figure 4-6J).

Knocking down of *ATR1* or *ATR2* did not affect plant growth. Transgenic *CYP81A12* had a lower fresh mass with a smaller height of the plants in comparison to Col_0 (Supporting Information Figure 4-1).

atr1 mutant increases herbicide sensitivity in *CYP81A12* line

Based on whole-plant dose response, Arabidopsis *atr1* and *atr2* mutant lines carrying *CYP81A12* treated with herbicides were more sensitive to all herbicides tested in the study, including penoxsulam, chlorsulfuron, mesotrione, and 2,4-D (Figure 4-7, Supporting Information Figures 4-2, 4-3). The transgenic Arabidopsis *CYP81A12* had a resistance index of 6.8, 23.6 and 20.9 for penoxsulam, mesotrione and 2,4-D, respectively (Table 4-5). When *ATR1* was knocked down (*CYP81A12 atr1-b*), the RI decreased to 1.0, 5.7, and 4.2, respectively to penoxsulam, mesotrione and 2,4-D. The chlorsulfuron doses used in this study were not able to control the transgenic Arabidopsis *CYP81A12* and *CYP81A12 atr1-a* (Figure 4-7B and Supporting Information Figures 4-3). The resistance to chlorsulfuron of these lines was very high. The lines *CYP81A12 atr2-a* and *CYP81A12 atr2-b* showed that the highest doses imparted a slight plant control and the model was fitted to the three-parameter model. These lines had a RI of 97.7 and 86.7 for chlorsulfuron (Table 4-5). The line *CYP81A12 atr1-b* was able to significantly reduce the RI to 3.7 (Table 4-5). *CYP81A12 atr1-b* had the highest decreasing in the RI to all herbicides. *ATR2* knockdown lines (*CYP81A12 atr2-a* and *CYP81A12 atr2-b*) also reduced the RI, but not with the same magnitude as *atr1-b* (Table 4-5).

The control of different herbicide doses on Arabidopsis lines grown in ½ MS medium yielded similar results (Figure 4-8 and Table 4-6). Among the mutants, the *CYP81A12 atr1-b*

line exhibited the most significant reversal of resistance to bensulfuron, penoxsulam, chlorsulfuron, propoxycarbazone, mesotrione, and bentazon, as evidenced by the highest magnitude of response (Figure 4-8). Notably, *CYP81A12 atr1-b* displayed symptoms similar to the Col_0 line at lower herbicide doses (Figure 4-8). In contrast, *CYP81A12 atr1-a*, *CYP81A12 atr2-a*, and *CYP81A12 atr2-b* did not exhibit reduced herbicide resistance compared to *CYP81A12*. Specifically, *CYP81A12* displayed resistance indices (RI) of 111.0, 35.7, 2.1, 4.2, 4.6, and 107.3 for bensulfuron, chlorsulfuron, penoxsulam, propoxycarbazone, mesotrione, and bentazon, respectively, while *CYP81A12 atr1-a* demonstrated reduced RIs of 22.3, 10, 0.9, 1.2, 0.9, and 0.9 for the same herbicides, reflecting similar GR_{50} values as Col_0 for penoxsulam, propoxycarbazone, mesotrione, and bentazon (Table 4-6).

Reduction of herbicide metabolism

Metabolism of mesotrione and chlorsulfuron was analyzed in the different Arabidopsis lines after herbicide application. The transgenic Arabidopsis line *CYP81A12* had a smaller peak area for the parental mesotrione than Col_0 in all time points after herbicide application, indicating faster herbicide metabolism (Figure 4-9A-B). The transgenic lines *CYP81A12 atr1-a*, *CYP81A12 atr2-a* and *CYP81A12 atr2-b* had similar profiles, with smaller peak area of parental mesotrione than Col_0. The only Arabidopsis line that showed the same peak area as Col_0 was *CYP81A12 atr1-b*. Hydroxylation of mesotrione was higher than Col_0 for all lines in all times points, except for *CYP81A12 atr1-b* (Figure 4-9A-B), which had a very small amount of hydroxy-mesotrione, highlighting that knockdown of *ATR1* prevents mesotrione metabolism in *CYP81A12 atr1-b*. Similar results were observed for chlorsulfuron metabolism (Figure 4-10A-B). Transgenic *CYP81A12* had a significantly smaller peak area of parental chlorsulfuron than Col_0, inferring faster herbicide metabolism. The transgenic mutant line *CYP81A12 atr1-a* was

not different in area peak of parental chlorsulfuron and hydroxy-chlorsulfuron when compared to Col_0, indicating a loss of the metabolism capacity.

NADPH-cytochrome P450 reductase in weeds

Thirteen different weed species were analyzed, and their corresponding CPRs (cytochrome P450 reductase) were identified (Figure 4-11). The number of CPR copies varied within each weed species, ranging from two to three copies (Supporting Information Table 4-2). Through the construction of a phylogenetic tree, it was observed that CPRs can be categorized into two distinct classes, namely CPR I and CPR II (Figure 4-11). These findings align with previous research, which suggests that CPR Class I is exclusive to dicot species, while Class II is present in both monocots and dicots[5, 27].

To differentiate between the CPRs within weed species, they were denoted as "a", "b" or "c" and individual copies were numbered as ".1", ".2" or ".3" to represent one, two, or three copies, respectively. Kochia (*Bassia scoparia*), common waterhemp, common morningglory (*Ipomea purpurea*), and Djulis (*Chenopodium formosanum*) have both CPR 1 and CPR 2 (Figure 4-11). Common morning glory possesses two copies of CPR 1 and one copy of CPR 2, whereas Djulis has three copies of CPR 1 and three copies of CPR 2. Kochia and common waterhemp possess one copy of each CPR class. On the other hand, Palmer amaranth and horseweed possess two copies of CPR 2, with horseweed having an additional third copy, while neither species has any copies of CPR 1. Class II Palmer amaranth CPRs have 93% identity, while the CPRs of horseweed (Class II) had a 74-76% identity among their three copies. Djulis, a hexaploid species with 27 chromosomes, exhibits three copies of each CPR class. Notably, each copy is localized within a different subgenome (B, C, and D).

Monocot plants exhibit diverse CPR compositions within CPR Class II, forming a distinct branch. Although their classification into specific classes is not documented in the literature, various monocot species possess CPR variants within Class II. For instance, blackgrass and goosegrass each possess three different CPR II variants: Am_CPRa, Am_CPRb, and Am_CPRc for blackgrass, and Ei_CPRa, Ei_CPRb, and Ei_CPRc for goosegrass. These CPR variants are categorized into separate branches based on phylogenetic analysis. Similarly, annual bluegrass, a tetraploid species, possesses three CPR variants, with two copies of each CPR variant present in subgenomes A or B. Similarly, late watergrass, an allotetraploid species, exhibits two copies of each of the three CPR variants, distributed across its subgenomes. Jungle rice, an allohexaploid species, possesses nine CPR variants, with copies of each CPR variant located in subgenomes D, E, or F. On the other hand, barnyardgrass, also an allohexaploid species, diverges from jungle rice as it possesses eight CPR variants. Among these variants, two CPRs have three copies each, while one CPR has two copies. It is noteworthy that downy brome is the sole monocot weed species analyzed in this study with only two CPR variants, whereas the remaining weed species all possess three distinct CPR variants with different copies numbers.

All CPRs had a well conserved domain (Supporting Information Figure 4-5). Among the identified CPRs in weed genomes, three CPRs, namely Eco_CPRc.1, Ec_CPRc.3, and Eo_CPRc.1, stand out with significantly larger protein sizes compared to the others. These three CPRs have 1520, 1905, and 1948 amino acids, respectively, whereas the remaining CPRs range from 689 to 823 amino acids. Eo_CPRc.1, Ec_CPRc.3 had a bigger N-terminal and additionally Eco_CPRc.1 have a long C-terminal when comparing to the other CPRs. Additionally, these three CPRs exhibit a similarity ranging from 87% to 94% among themselves. Additionally, there is close genetic relationship between subgenome D of Jungle rice and subgenome C of

Barnyardgrass[28, 29]. This suggests a common diploid ancestor, explaining the high similarity observed in CPRs among these three species.

DISCUSSION

Cytochrome P450 reductases (CPR) are a class of enzymes that play a crucial role in the electron transfer system from NADPH to P450s[5, 30]. While there are numerous P450 genes, the number of CPR genes is much smaller[31]. In Arabidopsis, two CPR, class I *ATR1* and class II *ATR2*, have been identified. *ATR1* is consistently expressed, while *ATR2* is induced in response to stress[5, 30]. Since P450s are implicated in the herbicide metabolic herbicide resistance in weeds, it is important to understand the specific role of P450 reductases in the herbicide-metabolizing capabilities of these P450s.

Through the generation of transgenic mutant Arabidopsis lines, i.e., transgenic Arabidopsis carrying *CYP81A12* from late watergrass able to metabolize a wide range of herbicides and with knockdown of either the *ATR1* or *ATR2* gene, our observations indicate that *ATR1* plays a predominant role in facilitating electron transfer from NADPH to *CYP81A12*. Upon knocking down the *ATR1* gene in the transgenic *CYP81A12* plants (*CYP81A12 atr1-b*), a 3.6-fold and 5.6-fold reduction in herbicide resistance for 2,4-D and mesotrione were observed, respectively. Conversely, the *CYP81A12 atr1-a* line, with a T-DNA insertion in the 5' untranslated region, did not exhibit a reduction in herbicide resistance due to insufficient gene expression reduction (Figure 4-6). Knockouts of the *ATR2* gene (*CYP81A12 atr2a* and *atr2-b*) also led to a decrease in herbicide resistance, although not to the same extent as *ATR1*.

In Arabidopsis, *ATR1* is commonly referred to as *CPR1*, while *ATR2* is known as *CPR2* in other species. Among the identified CPRs in weeds, it was observed that only CPR class II is present in grasses. In our current study, the *ATR1* of Arabidopsis was the most crucial in

providing electrons to *CYP81A12*. We hypothesize that different CPRs may exhibit varying interactions with various P450s, including those involved in herbicide metabolism. A recent investigation utilizing human CPRs indicated that mutations in the FMN binding domain significantly enhanced CYP activity. This suggests that amino acid variations within this domain play a significant role in the diverse interactions between CPRs and P450s[32]. Based on our results, it is evident that the presence of CPR in weed species exhibits variation. Notably, all examined grasses possess at least three different homologs of CPRs class II, except for downy brome, which has two. This discrepancy in the number of CPRs among grasses suggests duplication events that occurred prior to species differentiation within the grass family[5]. A detailed analysis of protein sequences revealed substantial differences primarily in the N- and C-terminal regions. However, the FMN-binding and FAD-binding domains in all CPRs displayed a high degree of conservation due to their specialized functions (Supporting Information Figure 4-5) and because of that, it is hypothesized that electrostatic interactions are considered to play a crucial role in guiding the formation of P450:CPR complexes.

The presence of varying numbers of CPRs in different weed species suggests distinct genetic pressures these weeds have encountered over time, leading to the acquisition of diverse functions. If herbicide application has selected resistant plants with enhanced herbicide-metabolizing activity, it becomes evident that it also selected weed biotypes with heightened CPR activity to support the increased activity of P450s by facilitating electron transfer; however, this aspect needs further investigation. Additionally, in our study, we observed that in Arabidopsis, *ATR2* is unable to complement the electron transfer to *CYP81A12* when *atr1* is knocked out. However, this may not be the case for other weed species that possess different CPR isoforms. For instance, jungle rice harbors three CPR isoforms, each with three copies.

These CPRs could potentially be paralogs with distinct functions, interacting differentially with various P450 enzyme families, or they may have complementary functions. Another relevant factor is that Arabidopsis *ATR1* and *ATR2* are linked, leading to the hypothesis that double knockout plants would be inviable. However, the absence of double wildtype plants in the F2 generation dismissed this hypothesis (Supporting Information Table 4-1). For most weed species examined in this study, the CPRs are dispersed across different chromosomes and not exhibit linkage. Nevertheless, our laboratory is actively pursuing the generation of a double knockout Arabidopsis line using CRISPR/Cas 9 to gain further insights into the role of *ATR1* and *ATR2* in metabolic resistance mechanisms in weeds.

Research on the role of CPR in the evolution of pesticide resistance has primarily focused on insects, with limited investigations into the functions of CPRs in plants. In insect studies, inhibiting CPR in deltamethrin-resistant *Cimex lectularius* increased susceptibility to deltamethrin, indicating that CPR could be targeted to reduce P450 activity[33]. Similarly, downregulating CPR in resistant *Tetranychus cinnabarinus* (carmine spider mite) and *T. urticae* (spider mite) through RNA interference reduced P450 activity and increased susceptibility to fenpropathrin and various acaricides, respectively[34]. In plants, which possess one, two, or three CPR isoforms, no research exists regarding the diverse functions of CPRs for herbicide evolution. For instance, Arabidopsis *ATR2*, which is co-expressed with lignin biosynthetic genes, showed significant alterations in lignin composition when mutated, leading to reduced activity of P450 enzymes such as cinnamate 4-hydroxylase, p-coumarate 3-hydroxylase, and ferulate 5-hydroxylase[35, 36]. The crystal structure of Arabidopsis *ATR2* revealed that mutations in the interflavin electron transfer region can impact ATR2 activity[37], offering opportunities to exploit these reliable enzymes for reducing P450 activity against herbicide metabolism.

Protein alignments revealed the presence of conserved domains shared across various CPRs, along with significant variation in the FMN binding domain (Supporting Information Figure 4-5). There were conserved regions among CPRs in both dicot and monocot weeds, presenting a valuable opportunity to target and reduce P450 activity. Inhibiting CPR activity, either through chemical inhibitors or RNA-targeting approaches, offers a potential means to restore herbicide effectiveness against weeds exhibiting metabolic herbicide resistance. Combining CPR inhibition with herbicide treatments can serve as a strategy to control metabolic herbicide-resistant weed populations and, additionally, delay the evolution of metabolic herbicide resistance in sensitive weed populations.

The study of cytochrome P450 reductases and their interactions with cytochrome P450 enzymes is crucial for understanding metabolic herbicide resistance in weeds. Our findings highlight the importance of CPRs, particularly *ATR1* in Arabidopsis, in transferring electrons to P450 involved in herbicide metabolism. The presence of varying numbers of CPRs in different weed species indicates the diverse genetic pressures these weeds have undergone. Inhibiting CPR activity could serve as a promising approach to mitigate herbicide resistance, control weed populations and prolong the activity of herbicides. Further research on CPRs and their functions in plants will contribute to our understanding of herbicide resistance evolution and the development of effective weed management strategies.

TABLES

Table 4-1. Primers and their characteristics used to genotype by PCR *ATR1* and *ATR2* genes with the T-DNA insertion and *CYP81A12* in F1 or F2 plants.

Gene	line	primer	sequence 5' to 3'	amplicon size (bp)
ATR1 (AT4G24520)	*atr1-a*	LP	ATCATCGGCAGCATAGTCATC	1045
		RP	AATGAATGGGCACAAAAACAG	
	atr1-b	LP	TGATTCTTTCCTGCAAACCAC	1098
		RP	TTATCCGAAGAAATCAAAGCG	
ATR2 (AT4G30210)	*atr2-a*	LP	ACCTGAACAACCCACTCAATG	1085
		RP	ATTGGTTGCATCGTTATGCTC	
	atr2-b	LP	TCACTCTTCTCGTAAGGCAC	1173
		RP	GCTTCATACTCCCGAGTCTG	
-	-	LBb1.3	ATTTTGCCGATTTCGGAAC	
CYP81A12	-	*Forward*	ATGGATAAGGCCTACGTGGCC	1557
	-	*Reverse*	TCAGAGCTCCTGAAGAACATCATG	

LP – left border, RP – right border, LBb1.3 - primer designed in the right border of the T-DNA sequence.

Table 4-2. Primers and their characteristics used for gene expression analysis of *ATR1*, *ATR2* and CYP81A12.

Gene	Exp.[1]	primer	Comment	sequence 5' to 3'	size (bp)
ATR1	qPCR	ATR1_A	*Forward* (spanning the *atr1-a* T-DNA insertion site)	CAAATCGGAAGCATACAAAG	167
			Reverse (spanning the *atr1-a* T-DNA insertion site)	CACCACCGATAACTCAATTA	
		ATR1_B	*Forward*	ATCATCCTCGTCCTTAGCCATA	172
			Reverse	TATGGGGACAGATTCGTTATCC	
		ATR1_C	*Forward* (spanning the *atr1-b* T-DNA insertion site)	TGTTCATATTGGCGATTACC	130
			Reverse (spanning the *atr1-b* T-DNA insertion site)	CCTACTGACAATGCTGCCA	
		ATR1_D	*Forward*	TCCTTCTGCACAGCAACATC	103
			Reverse	CTGAATACCGGGTGGTGACT	
ATR2	qPCR	ATR2_A	*Forward*	GTCCGTAGCTGCTGAATTATCC	103
			Reverse	GAGCATAACGATGCAACCAATA	
		ATR2_B	*Forward* (spanning the *atr2-a* T-DNA insertion site)	GAAGAAAGAGGATGTGGCTT	85
			Reverse (spanning the *atr2-a* T-DNA insertion site)	CCATTTGTAGAATCTCGCTGC	
		ATR2_C	*Forward*	TGGAATTTGACATTGCTGGA	167
			Reverse	GATTGGTGTGCCGTCTTCTT	
		ATR2_D	*Forward* (spanning the *atr2-b* T-DNA insertion site)	GTCAAAGAAGTCTACTTGAGGT	171
			Reverse (spanning the *atr2-b* T-DNA insertion site)	TCTCATAAACCAGTGCACATG	
		ATR2_E	*Forward*	TCCAGGGACTGGATTAGCTC	93
			Reverse	AACAAAACTGATGGCCCAAG	
ALS	qPCR and ddPCR	ALS_A	*Forward*	CATATGCTTGGAATGCATGG	99
			Reverse	CGTGACACGATCATCAAACC	
CYP81A12	qPCR and ddPCR	CYP	*Forward*	ACCTCCAGAGCATCATCCAC	136
			Reverse	TACACGTTCACCAGCAGCAT	

[1] Experiment, qPCR – primers were used for real-time qPCR gene expression analyses for the respective gene. ddPCR, primer used for genotyping the copy number of the respective gene by digital droplet PCR.

Table 4-3. PCR genotyping for T-DNA insertion in F2 segregating plants derived from crosses between *atr1* or *atr2* mutants with transgenic Arabidopsis *CYP81A12*.

Cross	Gene	Genotype	plants	frequency (%)
A	*ATR1*	HM-WT	7	12.1
		HM-M	15	25.8
		HZ	36	62.1
		total	58	
B	*ATR1*	HM-WT	16	25.8
		HM-M	12	19.3
		HZ	34	54.8
		total	62	
C	*ATR2*	HM-WT	20	34.8
		HM-M	9	15.5
		HZ	29	50
		total	58	
D	*ATR2*	HM-WT	9	17.3
		HM-M	15	28.8
		HZ	28	53.84
		total	52	

HM-WT – homozygous wild type, HM-M – homozygous mutant, HZ – heterozygous. Cross A – transgenic *CYP81A12* × *atr1-a*, cross B – transgenic *CYP81A12* × *atr1-b*, cross C – transgenic *CYP81A12* × *atr2-a*, and cross D – transgenic *CYP81A12* × *atr2-b*.

Table 4-4. ddPCR results of parental transgenic Arabidopsis *CYP81A12* and four selected F3 plants from crosses A (*CYP81A12* × *atr1-a*), B (*CYP81A12* × *atr1-b*), C (*CYP81A12* × *atr2-a*), and D (*CYP81A12* × *atr2-b*).

Cross	Samples	conc (copies/μL)						ALS/ CYP81
		ALS	Pmin[1]	Pmax[2]	*CYP81A12*	Pmin	Pmax	
Parental	*CYP81A12*	25.2	21.5	26.4	24	22.6	27.9	1.1
	CYP81A12	12.6	11.7	15.4	13.6	10.8	14.4	0.9
A	*CYP81 atr1-a*	86	81	92	77.6	72.7	82.5	1.1
	CYP81 atr1-a	110	104	116	104	99	110	1.1
	CYP81 atr1-a	135	128	141	120	114	126	1.1
	CYP81 atr1-a	116	110	122	109	103	115	1.1
B	*CYP81 atr1-b*	72.7	68.1	77.4	72.1	67.4	76.8	1.0
	CYP81 atr1-b	83.1	78.1	88	74	69	79	1.1
	CYP81 atr1-b	50.8	47	54.7	49.7	45.7	53.6	1.0
	CYP81 atr1-b	78.6	73.8	83.4	75	70	80	1.0
C	*CYP81 atr2-a*	187	194	208	200	179	194	0.9
	CYP81 atr2-a	115	109	120	111	105	117	1.0
	CYP81 atr2-a	67.2	62.8	71.6	59.8	55.4	64.2	1.1
	CYP81 atr2-a	178	170	185	176	168	183	1.0
D	*CYP81 atr2-b*	45.1	41.3	48.8	45.7	41.8	49.6	1.0
	CYP81 atr2-b	86	81	91	89	84	94	1.0
	CYP81 atr2-b	77	72	82	75.2	70.4	80.2	1.0
	CYP81 atr2-b	98	93	104	89	84	94	1.1

[1]Pmin – Minimum target concentration normalized for the low error bar of the droplet Poisson distribution for the 95% confidence interval. [2]Pmax – Maximum target concentration normalized for the high error bar of the droplet Poisson distribution for the 95% Confidence Interval.

Table 4-5. Parameters of the logistic equation and resistance index (RI) for shoot fresh weight (% of untreated control) of transgenic Arabidopsis *CYP81A12* and *atr1*, *atr2* mutants treated with different doses of 2,4-D and mesotrione.

Line	Herbicide	b	GR50				RI	*p*-value
			Dose (g ha^{-1})	Lower CI (95%)	Upper CI (95%)	SE		
Col_0	PENO	0.5	2.3	0.9	3.7	0.7	-	
CYP81A12		0.7	15.7	6.7	24.6	4.5	6.8	<0.001
CYP81A12 atr1-a		1.8	17.1	11.2	23.1	3.0	7.4	0.01
CYP81A12 atr1-b		0.7	2.3	1.2	3.5	0.6	1.0	0.9
CYP81A12 atr2-a		2.9	6.7	5.0	8.4	0.8	2.9	0.04
CYP81A12 atr2-b		0.7	3.4	1.5	5.2	0.9	1.5	0.4
Col_0	CHLO	1.3	0.9	0.7	1.1	0.1	-	-
CYP81A12		-	-	-	-	-	-	-
CYP81A12 atr1-a		-	-	-	-	-	-	-
CYP81A12 atr1-b		2.4	3.4	2.5	4.4	0.4	3.7	<0.01
CYP81A12 atr2-a		2.6	87.9	1.8	174.1	42.5	97.7	0.03
CYP81A12 atr2-b		0.6	77.5	1.1	153.9	37.7	86.7	0.04
Col_0	MESO	3.7	4.4	3.85	5.1	0.3	-	-
CYP81A12		15	93.6	58.6	128.6	17.7	20.9	<0.001
CYP81A12 atr1-a		5.5	89.3	78.9	99.7	5.2	19.9	<0.001
CYP81A12 atr1-b		4.3	26.8	23	30.5	1.9	5.7	<0.001
CYP81A12 atr2-a		2.8	41.8	35.5	48.1	3.2	9.3	<0.001
CYP81A12 atr2-b		2.9	59.1	50.3	68.1	4.5	13.2	<0.001
Col_0	2,4-D	0.5	5.9	2.7	9.3	1.6	-	-
CYP81A12		1.2	141.7	96.5	186.5	22.7	23.6	<0.001
CYP81A12 atr1-a		1.6	103.3	73.7	132.8	14.9	17.2	<0.001
CYP81A12 atr1-b		2.9	25.3	20.15	30.4	2.6	4.2	<0.001
CYP81A12 atr2-a		0.8	50.1	29.6	70.5	10.3	8.3	<0.001
CYP81A12 atr2-b		3.8	130.1	108.1	151.9	11.1	21.6	<0.001

b: slope; d: upper limit was fixed to 100; GR50: herbicide dose that causes a 50% reduction in shoot fresh weight variable; CI: confidence interval of parameter GR50 ($\alpha = 0.05$); RI: resistance index = GR50 ratio between lines with Col_0 (control). PENO – penoxsulam, CHLO – chlorsulfuron, MESO – mesotrione.

Table 4-6. Parameters of the logistic equation and resistance index (RI) for leaf number of Arabidopsis seedlings (% of untreated control) of wild-type, *atr1-a*, *atr1-b*, *atr2-a*, *atr2-b,* and Arabidopsis transgenic *CYP81A12*, *CYP81A12 atr-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a*, *CYP81A12 atr2-b* growing in ½ MS medium supplemented with different doses of herbicides.

Lines		b	ED50				RI	*p*-value
			dose (nM)	Lower CI (95%)	Upper CI (95%)	SE		
Col_0	BSM	35.7	9.5	9.1	9.8	0.17	-	-
atr1-a		32.9	3.4	-16.3	23.0	10.0	0.3	0.5
atr1-b		31.5	9.0	8.2	9.9	0.4	0.9	0.4
atr2-a		27.7	9.1	8.3	9.8	0.4	0.9	0.4
atr2-b		26.6	6.7	-12.9	26.4	9.9	0.8	0.8
CYP81A12		18.3	1055.5	1034.3	1071.8	9.5	111.0	<0.001
CYP81A12 atr1-a		3.6	1368.4	1321.6	1360.9	9.9	141.4	<0.001
CYP81A12 atr1-b		24.9	185.2	192.4	231.7	10.0	22.3	<0.001
CYP81A12 atr2-a		2.0	1547.8	1528.2	1567.5	9.9	163.2	<0.001
CYP81A12 atr2-b		5.6	942.6	923.1	962.2	9.9	99.4	<0.001
Col_0	CHL	3.6	3.7	0.73	6.73	1.5	-	-
atr1-a		1.3	1.2	0.8	1.5	0.2	0.3	<0.01
atr1-b		1.7	2.4	1.4	3.3	0.5	0.6	0.03
atr2-a		7.3	3.6	-1.1	8.4	2.4	0.9	0.19
atr2-b		2.1	1.9	1.4	2.3	0.2	0.5	<0.01
CYP81A12		1.2	132.2	81.9	182.6	25.08	35.7	0.02
CYP81A12 atr1-a		2.7	121.6	93.9	149.2	13.8	32.8	<0.01
CYP81A12 atr1-b		2.3	37.2	27.74	46.66	4.7	10.0	<0.01
CYP81A12 atr2-a		5.4	129.9	8.5	251.4	60.5	35.1	0.1
CYP81A12 atr2-b		0.7	76.1	17.5	134.6	29.1	20.5	0.07
Col_0	PEN	3.6	13.8	11.8	15.7	0.95	-	-
atr1-a		1.4	3.8	2.6	4.9	0.5	0.3	<0.001
atr1-b		0.9	3.3	0.5	6.1	1.4	0.2	0.02
atr2-a		1.0	2.7	1.1	4.37	0.8	0.2	<0.001
atr2-b		3.1	7.0	5.9	8.1	0.5	0.5	<0.001
CYP81A12		8.8	29.6	13.4	45.8	8.0	2.1	<0.001
CYP81A12 atr1-a		10.3	26.9	10.3	43.5	8.21	1.9	<0.01
CYP81A12 atr1-b		11.6	11.9	-8.8	32.8	10.4	0.9	0.25
CYP81A12 atr2-a		10.1	27.8	9.8	45.8	8.9	2.0	<0.01
CYP81A12 atr2-b		8.5	23.7	13.2	34.3	5.2	1.7	<0.001
Col_0	PRO	2.4	3.5	2.8	4.0	0.3	0.9	-
atr1-a		2.1	3.4	1.8	5.1	0.8	0.9	0.9
atr1-b		3.1	3.2	2.5	3.9	0.3	0.9	0.7
atr2-a		8.1	3.8	0.2	7.4	1.8	1.1	0.9
atr2-b		2.4	2.4	1.6	3.2	0.4	0.8	0.2

CYP81A12		2.4	15.5	12.1	19.0	1.7	4.2	<0.001
CYP81A12 atr1-a		1.9	13.0	10.9	15.1	1.0	3.7	<0.001
CYP81A12 atr1-b		2.1	4.3	3.7	4.9	0.3	1.2	0.2
CYP81A12 atr2-a		2.4	15.7	10.3	21.1	2.6	4.3	<0.001
CYP81A12 atr2-b		1.8	11.1	7.6	14.5	1.7	3.2	<0.001
Col_0		4.5	159.2	131.8	186.7	13.9	-	-
atr1-a		5.4	129.3	106.8	151.8	11.9	0.8	0.3
atr1-b		6.5	160.4	114.8	206.1	23.2	1	0.9
atr2-a		6.3	124.3	84.97	163.6	20.1	0.8	0.1
atr2-b		3.8	117.7	100.4	135.1	8.3	0.7	0.03
CYP81A12	MES	7.9	740.7	604.7	876.78	69.19	4.6	<0.0001
CYP81A12 atr1-a		2.9	480.2	415.1	545.2	33.1	3	<0.0001
CYP81A12 atr1-b		5.1	142.8	112.9	172.72	15.2	0.9	0.4
CYP81A12 atr2-a		6.1	564	492.2	634	36.1	3.5	<0.001
CYP81A12 atr2-b		2.9	585.2	495.7	675.1	45.6	3.6	<0.0001
Col_0		5.4	0.1	0.09	0.11	0.004	-	-
atr1-a		3.9	0.01	0.01	0.02	0.001	0.1	<0.001
atr1-b		3.1	0.03	0.02	0.05	0.007	0.3	<0.001
atr2-a		2.3	0.03	0.02	0.04	0.005	0.3	<0.001
atr2-b		2.2	0.02	0.01	0.03	0.002	0.2	<0.001
CYP81A12	BEN	5	11.24	9.57	12.9	0.82	107.3	<0.001
CYP81A12 atr1-a		3.7	12.84	10.7	14.9	1	132.3	<0.001
CYP81A12 atr1-b		5.22	0.08	-0.07	0.2	0.07	0.9	0.8
CYP81A12 atr2-a		8.43	11.7	3	20.5	4.35	111.6	<0.001
CYP81A12 atr2-b		3.4	12.2	10.2	14.3	1.01	114.6	<0.001

b: slope; d: upper limit was fixed to 100; GR50: herbicide dose that causes a 50% reduction in shoot fresh weight variable; CI: confidence interval of parameter GR50 ($\alpha = 0.05$); RI: resistance index = GR50 ratio between lines with Col_0 (control).

FIGURES

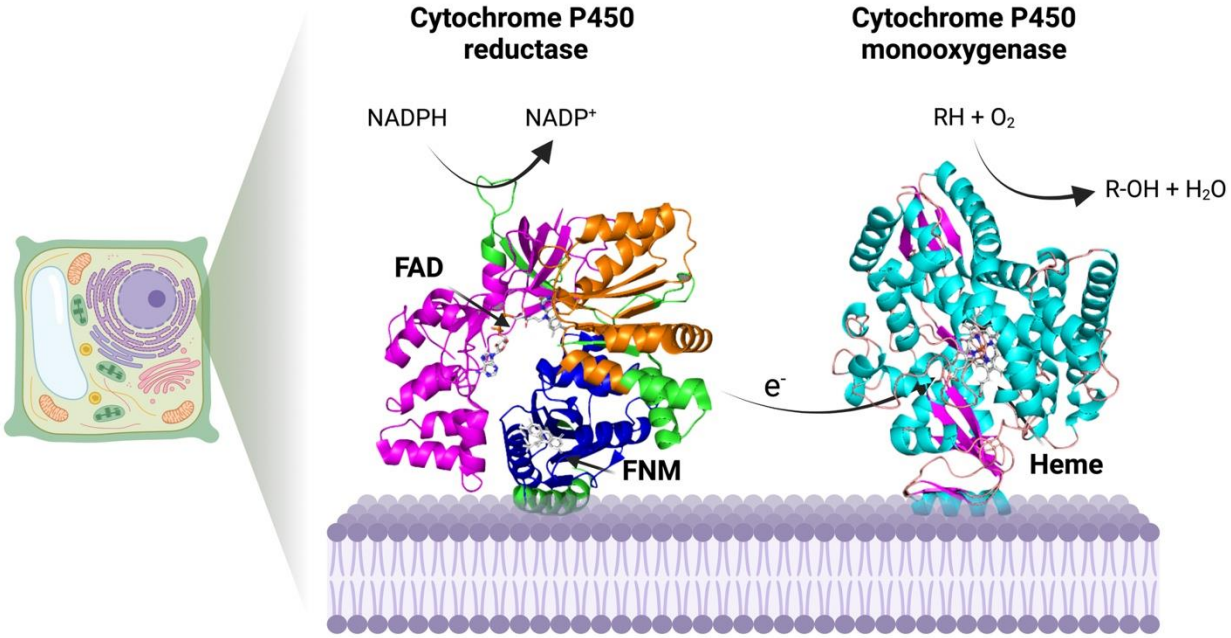


Figure 4-1. Process of electron transfer from NADPH-cytochrome P450 reductase to the catalytic site of cytochrome P450 monooxygenase. Two electrons are sequentially transferred from NADPH to the heme group of P450 through the cofactors FAD and FMN, each transferred in single-electron steps. Cytochrome P450, once activated by the electron transfer, catalyzes the oxidative breakdown of a substrate by inserting one oxygen atom into the chemical compound. Created with Biorender.com.

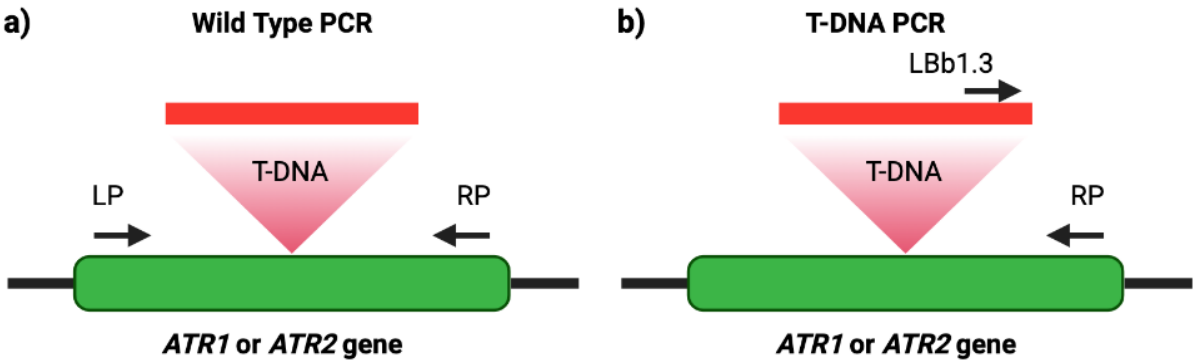


Figure 4-2. Two-step PCR for genotyping mutant lines. The "wild-type PCR" reaction will amplify a region that is present in wild type and heterozygous lines but will not amplify in homozygous lines. The "T-DNA PCR" will amplify only if T-DNA insertion is present. LP – left primer, RP – right primer, LBb1.3 – left border primer for T-DNA insertion.

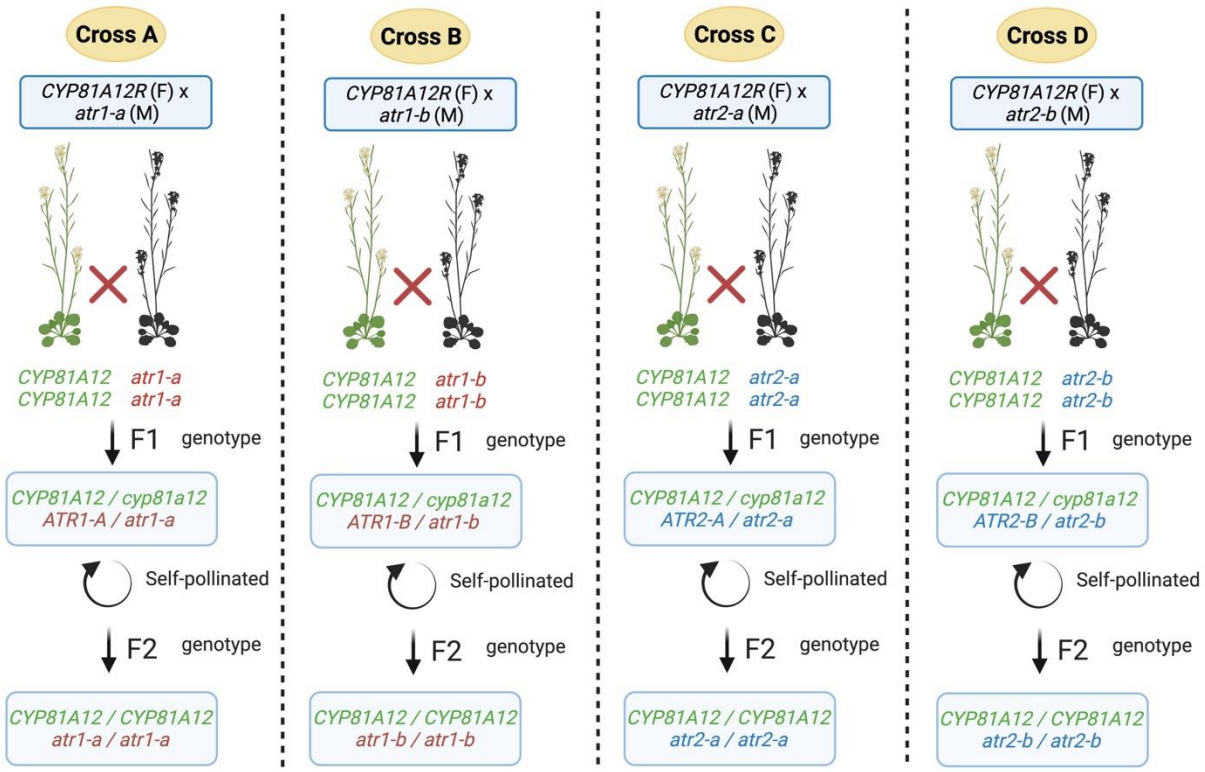


Figure 4-3. Schemes representing the Arabidopsis crosses to generate transgenic Arabidopsis *CYP81A12* with loss of function in *atr1* or *atr2*. F1 plants were genotyped for T-DNA insertion and *CYP81A12* presence. F2 plants were genotyped for homozygous T-DNA insertion and two copies of *CYP81A12* using ddPCR. M - male, F - female. Created with BioRender.com.

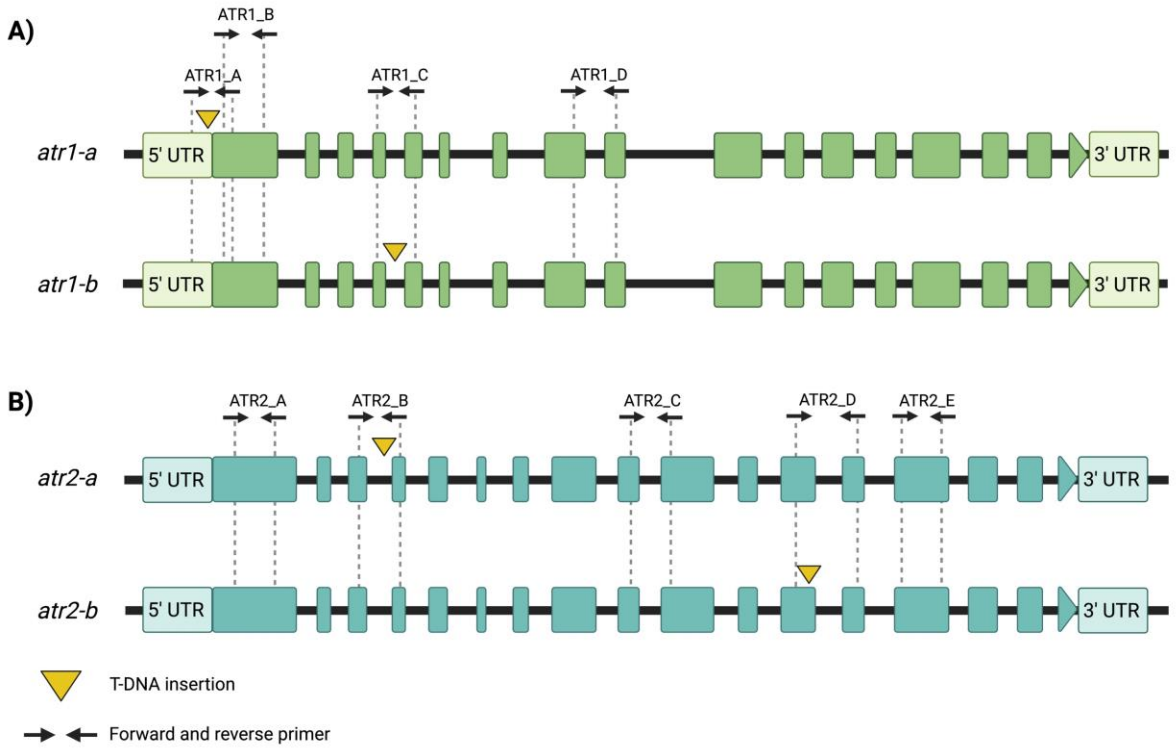


Figure 4-4. T-DNA location in *ATR1* (A) and *ATR2* (B) genes in SALK lines and primers localization for real time qPCR gene expression analysis. Yellow triangle locates the T-DNA insertion in the genes, *atr1-a* and *atr2-b* have T-DNA insertion in the 5'UTR and 4th intron, respectively, and *atr2-a* and *atr2-b* have T-DNA insertion in the 3rd and 12th exon, respectively. Boxes indicate exons or untranslated regions (UTR), black line indicates introns.

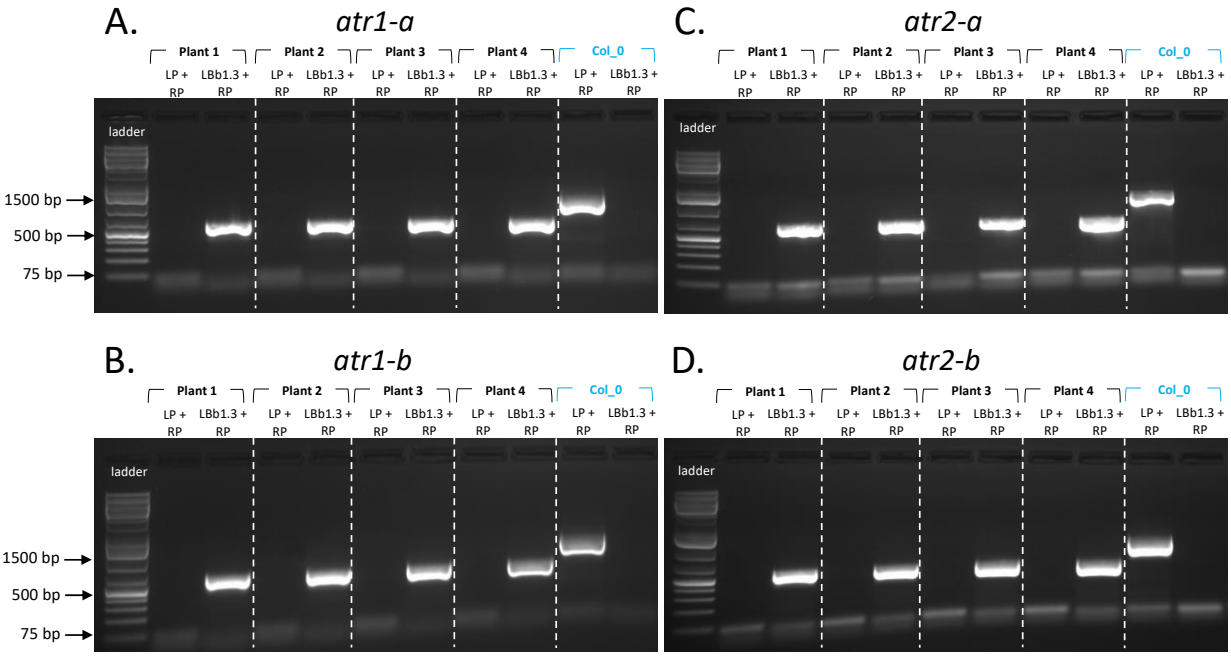


Figure 4-5. T-DNA insertion genotyping for *ATR1* and *ATR2* gene in different Arabidopsis mutant lines. Figures include four plants of each line and one wildtype columbia_0 (Col_0) as a control. Each plant was submitted for two PCR, first reaction was "wildtype PCR" using primer set LP + RP, and second reaction "T-DNA PCR" reaction using primers set LBb1.3 + RP.

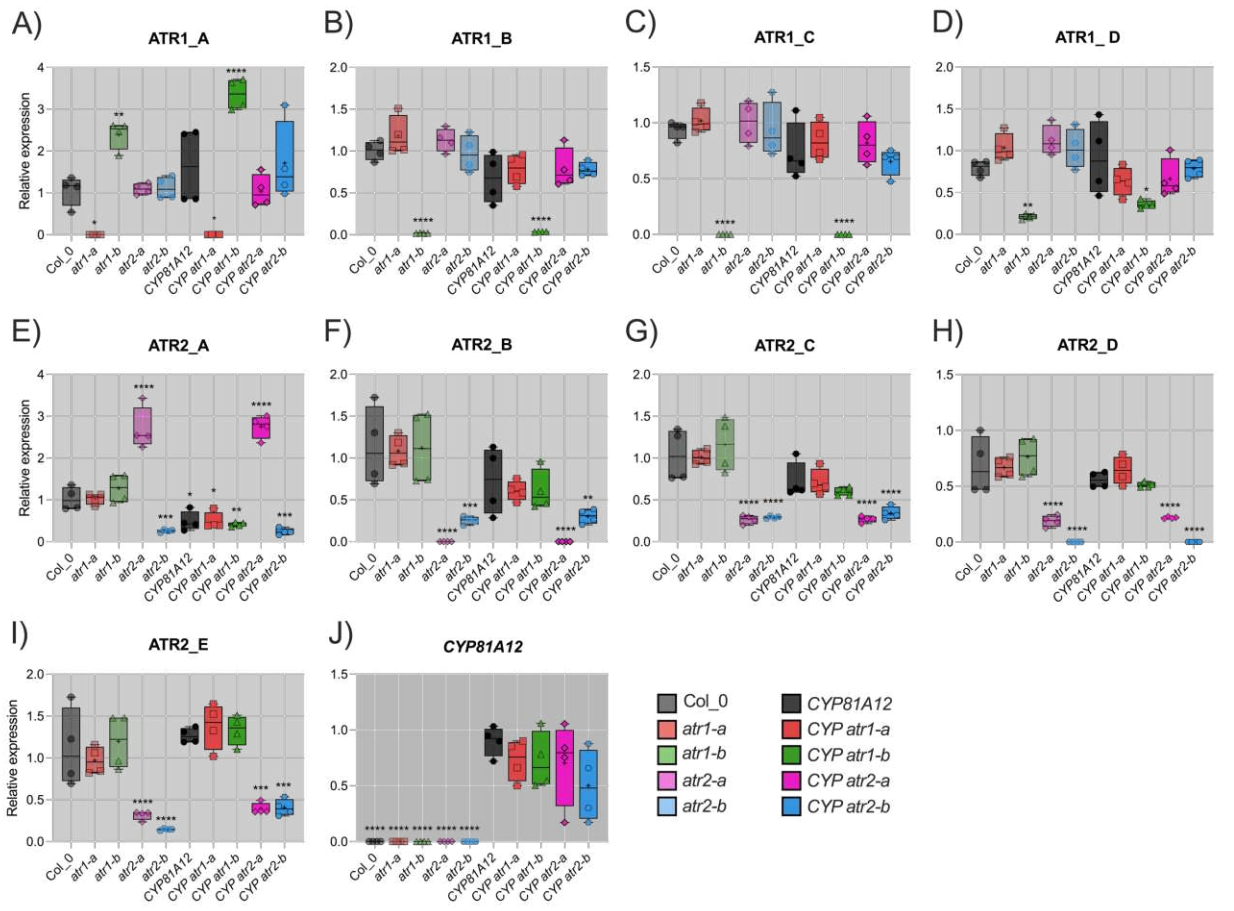


Figure 4-6. Relative gene expression of *ATR1* and *ATR2* gene in different mutant and transgenic mutant Arabidopsis lines. A-D) *ATR1* gene expression was analyzed in four different locations using primer set ATR1_A, ATR1_B, ATR1_C and ATR1_D. E-I) *ATR2* gene expression was analyzed in five different locations using primer set ATR2_A, ATR2_B, ATR2_C, ATR2_D and ATR2_E. J) Expression of *CYP81A12* with only one primer set. The primers set localization is illustrated in Figure 4-4. ATR1 and ATR2 gene expression was calculated using Col_0 as control. *CYP81A12* gene expression was calculated using transgenic Arabidopsis CYP81A12 as control. *Significant by Dunnett's Test.

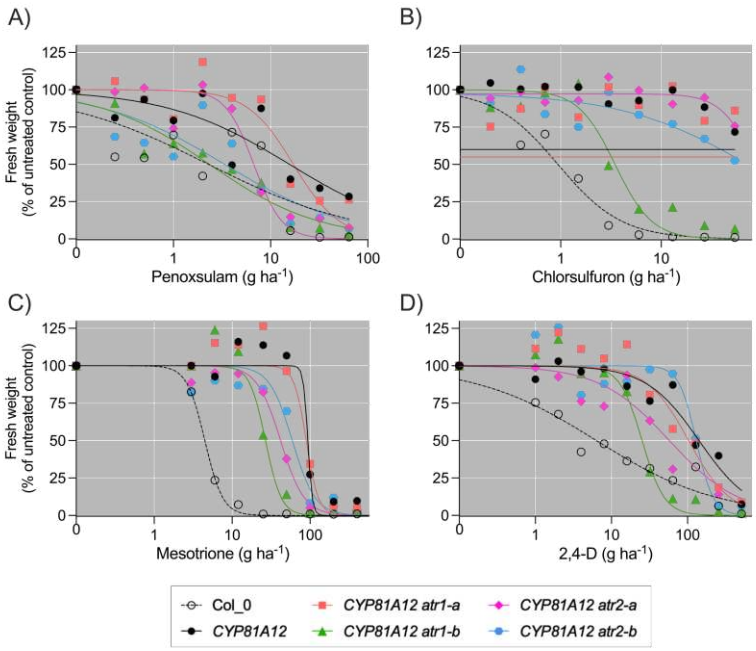


Figure 4-7. Dose-response of Arabidopsis wild-type (Col_0), transgenic (*CYP81A12*) and transgenic mutant lines *CYP81A12 atr1-a, CYP81A12 atr1-b, CYP81A12 atr2-a, CYP81A12 atr2-b.* Response of Arabidopsis lines to increasing doses of A) penoxsulam, B) chlorsulfuron, C) mesotrione and D) 2,4-D.

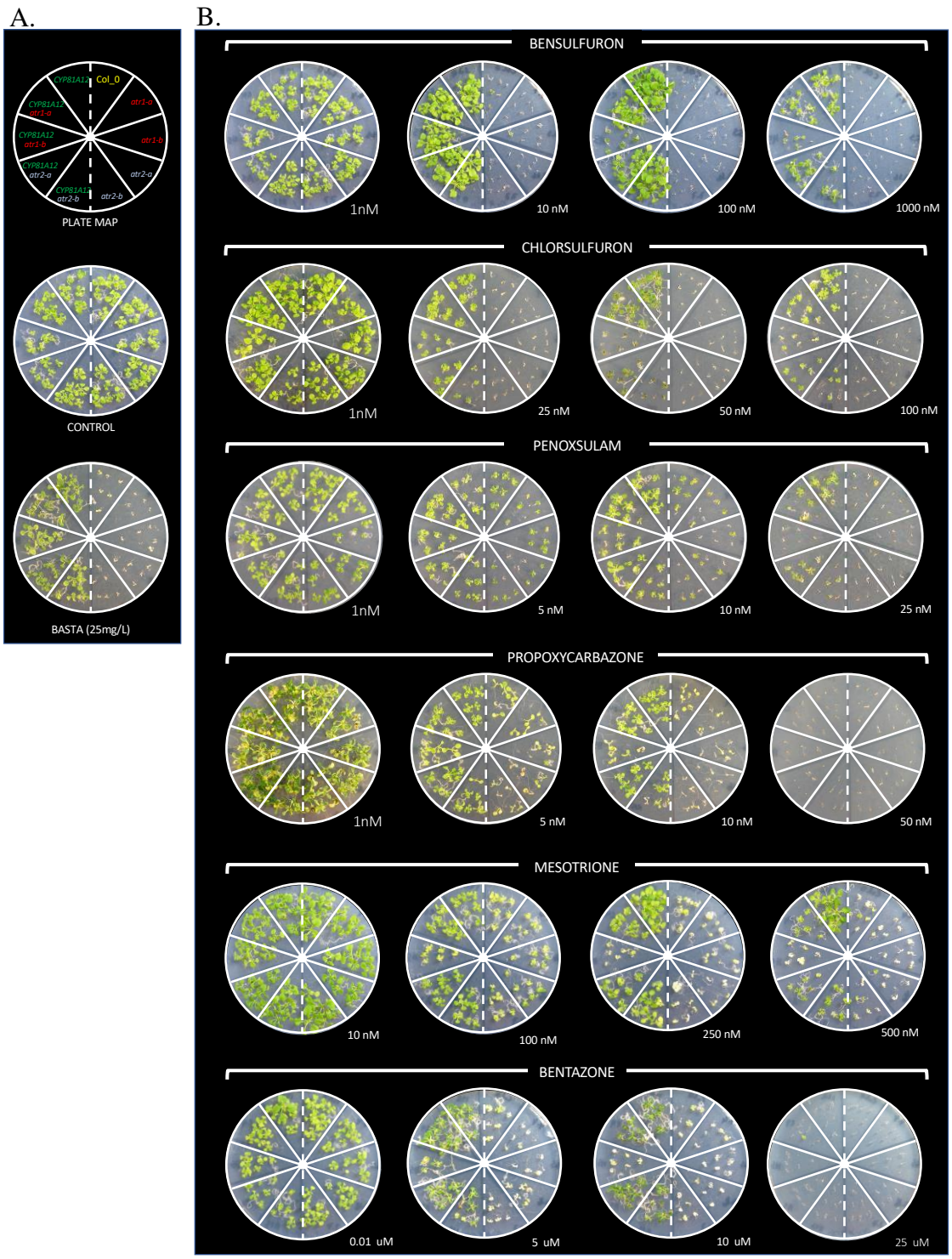


Figure 4-8. Sensitivity of wild type (Col_0), mutant *atr1-a*, *atr1-b*, *atr2-a*, *atr2-b*, transgenic Arabidopsis *CYP81A12* and transgenic mutant Arabidopsis *CYP81A12 atr1-a, CYP81A12 atr1-b*, *CYP81A12 atr2-a, CYP81A12 atr2-b*. Pictures at 14 d-old seedlings growing in ½ MS media containing different herbicide concentrations. A) Plate map, ½ MS control untreated and basta (25mg L^{-1}); B) ½ MS plates with different doses of bensulfuron-methyl, chlorsulfuron, penoxsulam, propoxycarbazone, mesotrione, and bentazon.

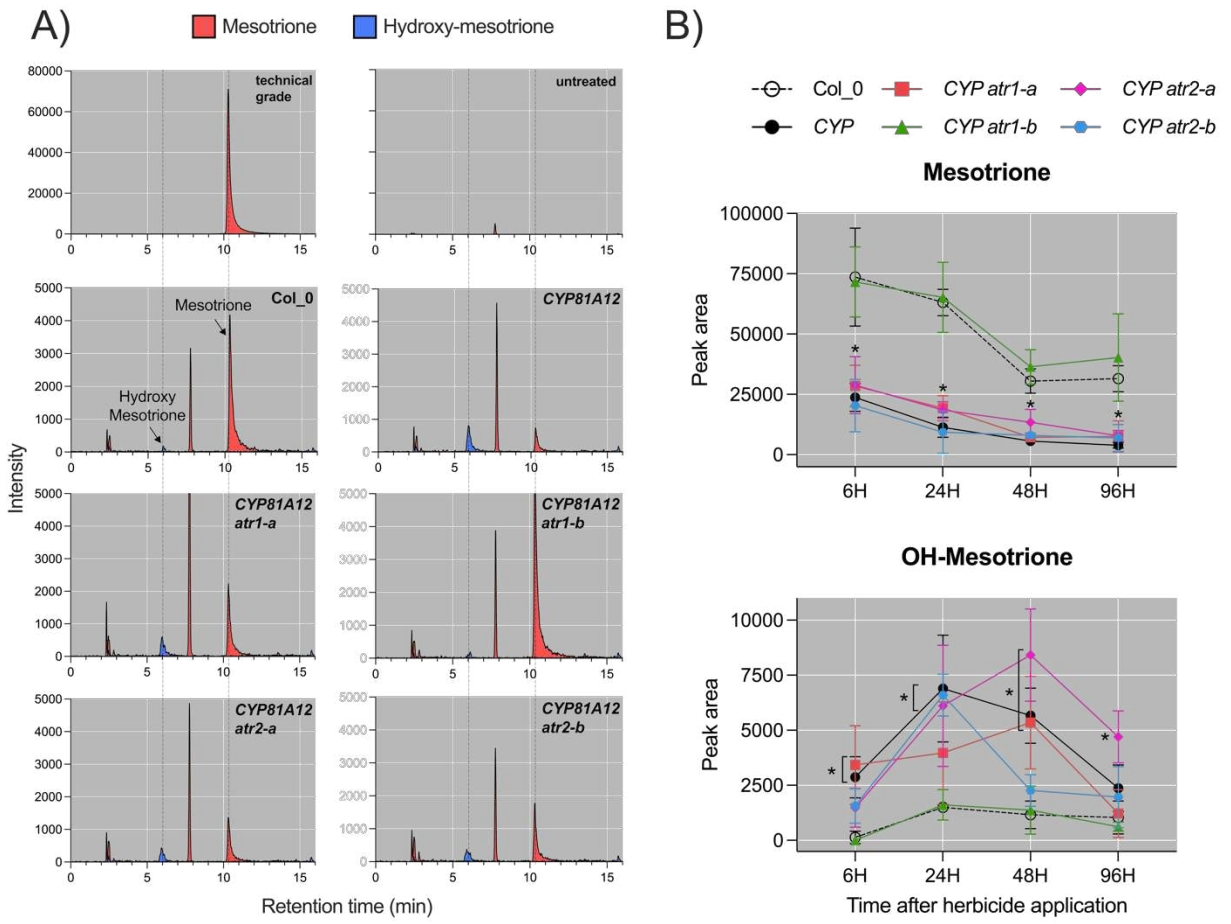


Figure 4-9. LC-MS/MS analysis of mesotrione metabolites found in different Arabidopsis lines. Arabidopsis lines analyzed were wild-type (Col_0), transgenic *CYP81A12* and transgenic mutants *CYP81A12 atr1-a*, *CYP81A12 atr1-b, CYP81A12 atr2-a, CYP81A12 atr2-b.* A) MS chromatogram 24 h after herbicide application. The retention time for parental mesotrione and hydroxy-mesotrione is 10.3- and 5.9-min, respectively. Peak at retention time 7.7 min is present in all samples, even in untreated plants. B) Peak area of mesotrione and hydroxy-mesotrione at 6, 24, 48 and 96 h after mesotrione application on different Arabidopsis lines. * *p-value* <0.05 by Dunnett's Test.

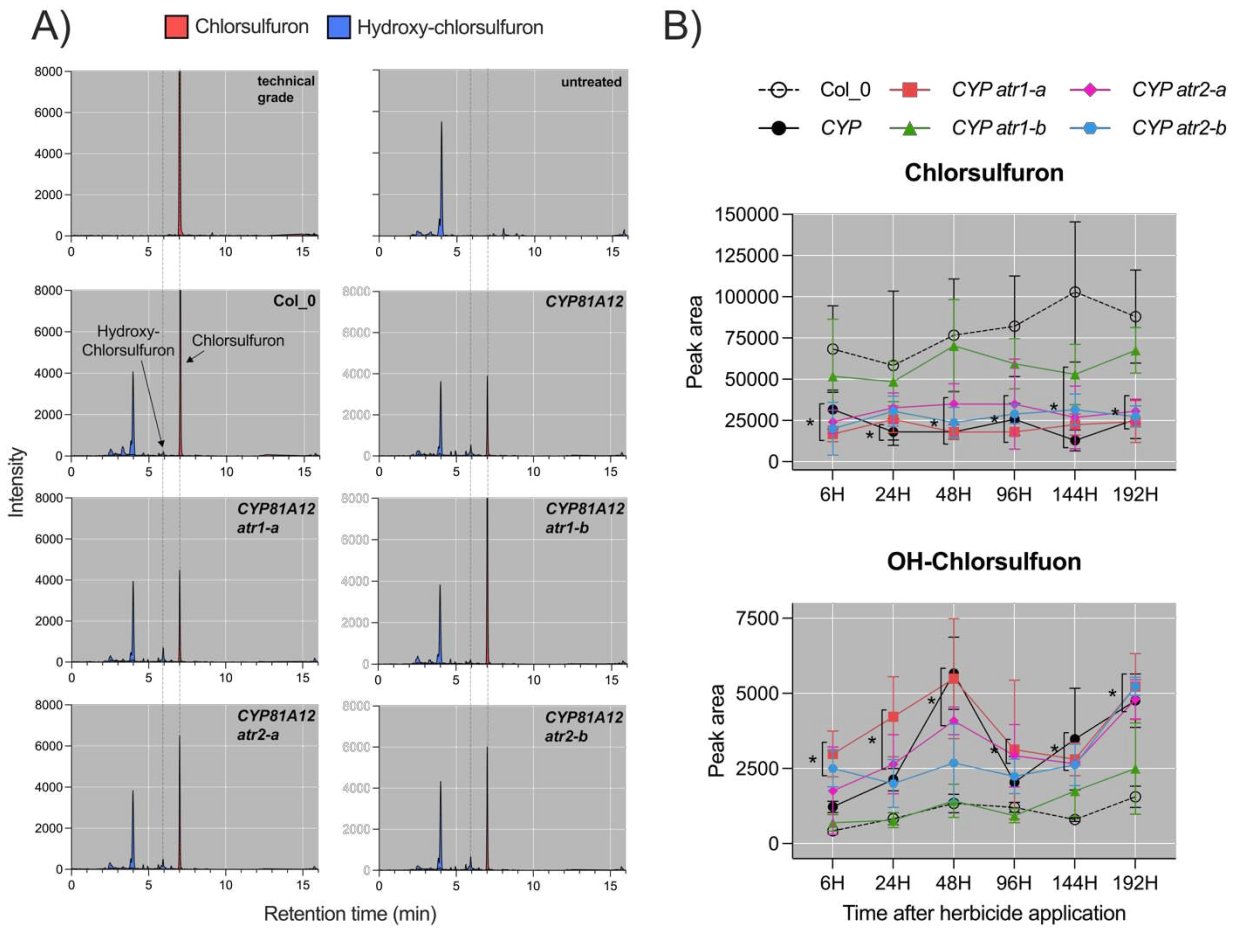


Figure 4-10. LC-MS/MS analysis of chlorsulfuron metabolites found in different Arabidopsis lines. Arabidopsis lines analyzed were wild-type (Col_0), transgenic *CYP81A12* and transgenic mutants *CYP81A12 atr1-a*, *CYP81A12 atr1-b, CYP81A12 atr2-a, CYP81A12 atr2-b.* A) MS chromatogram 192 h after herbicide application. The retention time for parental chlorsulfuron and hydroxy-chlorsulfuron 7.01 and 5.9 min, respectively. Peak at retention time 4.01 min is present in all samples, even in untreated plants. B) Peak area of chlorsulfuron and hydroxy-chlorsulfuron at different time points after chlorsulfuron application. * *p-value* <0.05 by Dunnett's Test.

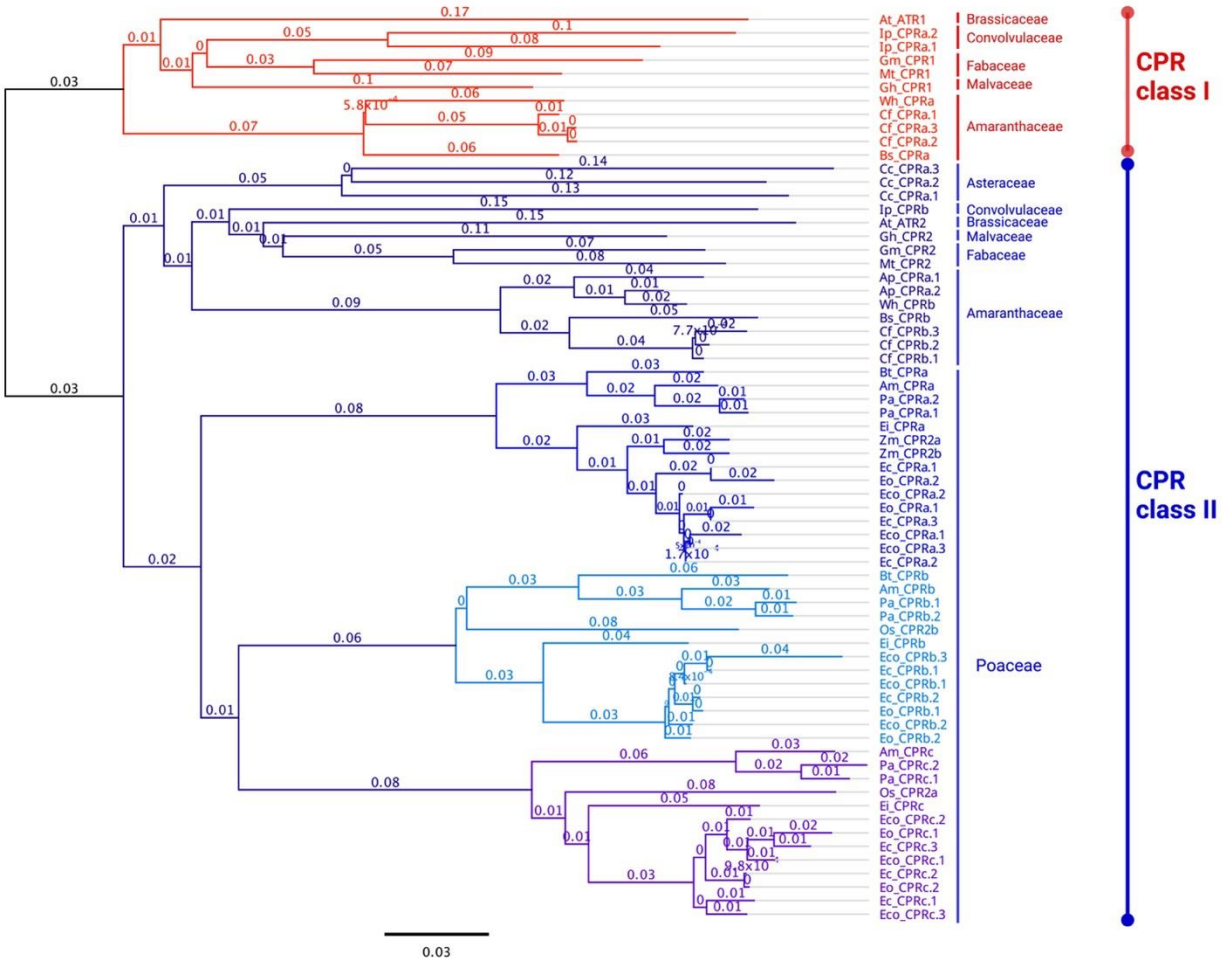


Figure 4-11. Phylogenetic tree of 67 CPR protein sequences from 13 different weed species and others 6 plant species with know CPRs sequences and classification. The phylogenetic tree shows branching of class I (red) and II (blue) CPRs. Different CPR were named as "a", "b" or "c" and the copies were numbered as ".1", ".2" or ".3" for one, two or three copies, respectively. CPR, cytochrome P450 reductase; Am - *Alopecurus myosuroides* (blackgrass), At - *Arabidopsis thaliana* (Arabidopsis), Ap – *Amaranthus palmeri* (Palmer amaranth), Wh – *Amaranthus tuberculatus* (waterhemp), Bs - *Bassia scoparia* (kochia), Bt - *Bromus tectorum* (downy brome), Cc - *Conyza canadensis* (horseweed), Cf - *Chenopodium formosanum* (Djulis), Ec - *Echinochloa crus-galli* (barnyardgrass), *Eo - E. oryzicola* (late watergrass), Eco - *E. colona* (jungle rice), Ei - *Eleusine indica* (goosegrass), Gh - *Gossypium hirsutum* (cotton), Gm - *Glycine max* (soybean), Ip - *Ipomoea purpurea* (common morning-glory), Mt - *Medicago truncatula* (Barrel medic), Os - *Oryza sativa* (rice), Pa - *Poa annua* (annual bluegrass).

REFERENCES

1. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A and Dayan FE, Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry* **295**: 10307–10330 (2020).
2. Rigon CAG, Gaines TA, Küpper A and Dayan FE, Metabolism-based herbicide resistance, the major threat among the non-target site resistance mechanisms. *Outlooks on Pest Manag* **31**: 162-168 (2020).
3. Pandian BA, Sathishraj R, Djanaguiraman M, Prasad PVV and Jugulam M, Role of cytochrome P450 enzymes in plant stress response. *Antioxidants* **9**: 454 (2020).
4. Wang M, Roberts DL, Paschke R, Shea TM, Masters BSS and Kim J-JP, Three-dimensional structure of NADPH–cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. *Proceedings of the National Academy of Sciences* **94**: 8411-8416 (1997).
5. Jensen K and Møller BL, Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* **71**: 132-141 (2010).
6. Porter TD and Kasper CB, NADPH-cytochrome P-450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. *Biochemistry* **25**: 1682-1687 (1986).
7. Paine MJI, Scrutton NS, Munro AW, Gutierrez A, Roberts GCK and Wolf CR. Electron transfer partners of cytochrome. In *Cytochrome P450, structure, mechanism, and biochemistry*, ed. by New York NS: New York (2005).
8. Lau SM and O'Keefe DP, Analysis of herbicide metabolism by monocot microsomal cytochrome P450. *Methods Enzymol* **272**: 235-242 (1996).
9. Munro AW, Noble MA, Robledo L, Daff SN and Chapman SK, Determination of the redox properties of human NADPH-cytochrome P450 reductase. *Biochemistry* **40**: 1956-1963 (2001).
10. Mukherjee G, Nandekar PP and Wade RC, An electron transfer competent structural ensemble of membrane-bound cytochrome P450 1A1 and cytochrome P450 oxidoreductase. *Communications Biology* **4**: 55 (2021).
11. Paquette SM, Jensen K and Bak S, A web-based resource for the Arabidopsis P450, cytochromes b5, NADPH-cytochrome P450 reductases, and family 1 glycosyltransferases (http://www.P450.kvl.dk). *Phytochemistry* **70**: 1940-1947 (2009).
12. Nelson DR, Schuler MA, Paquette SM, Werck-Reichhart D and Bak S, Comparative genomics of rice and Arabidopsis. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. *Plant Physiol* **135**: 756-772 (2004).
13. Wu D, Shen E, Jiang B, Feng Y, Tang W, Lao S, Jia L, Lin H-Y, Xie L, Weng X, Dong C, Qian Q, Lin F, Xu H, Lu H, Cutti L, Chen H, Deng S, Guo L, Chuah T-S, Song B-K, Scarabel L, Qiu J, Zhu Q-H, Yu Q, Timko MP, Yamaguchi H, Merotto A, Qiu Y, Olsen KM, Fan L and Ye C-Y, Genomic insights into the evolution of Echinochloa species as weed and orphan crop. *Nat Commun* **13**: 689 (2022).
14. Pan G, Zhang XY, Liu KD, Zhang JW, Wu XZ, Zhu J and Tu JM, Map-based cloning of a novel rice cytochrome P450 gene *CYP81A6* that confers resistance to two different classes of herbicides. *Plant MolBiol* **61**: 933-943 (2006).
15. Peng YH, Lai Z, Lane T, Nageswara-Rao M, Okada M, Jasieniuk M, O'Geen H, Kim RW, Sammons RD, Rieseberg LH and Stewart CN, De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms. *Plant Physiol* **166**: 1241-1254 (2014).
16. Urban P, Mignotte C, Kazmaier M, Delorme F and Pompon D, Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 *CYP73A5*. *Journal of Biological Chemistry* **272**: 19176-19186 (1997).

17. Mizutani M and Ohta D, Two isoforms of NADPH:cytochrome P450 reductase in Arabidopsis thaliana. Gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol* **116**: 357-367 (1998).
18. Varadarajan J, Guilleminot J, Saint-Jore-Dupas C, Piégu B, Chabouté ME, Gomord V, Coolbaugh RC, Devic M and Delorme V, ATR3 encodes a diflavin reductase essential for Arabidopsis embryo development. *New Phytol* **187**: 67-82 (2010).
19. Rana S, Lattoo SK, Dhar N, Razdan S, Bhat WW, Dhar RS and Vishwakarma R, NADPH-cytochrome P450 reductase: molecular cloning and functional characterization of two paralogs from *Withania somnifera* (L.) Dunal. *PLoS One* **8**: 16 (2013).
20. Parage C, Foureau E, Kellner F, Burlat V, Mahroug S, Lanoue A, Dugé de Bernonville T, Londono MA, Carqueijeiro I, Oudin A, Besseau S, Papon N, Glévarec G, Atehortùa L, Giglioli-Guivarc'h N, St-Pierre B, Clastre M, O'Connor SE and Courdavault V, Class II cytochrome P450 reductase governs the biosynthesis of alkaloids. *Plant Physiol* **172**: 1563-1577 (2016).
21. Dimaano NG, Yamaguchi T, Fukunishi K, Tominaga T and Iwakami S, Functional characterization of cytochrome P450 *CYP81A* subfamily to disclose the pattern of cross-resistance in *Echinochloa phyllopogon*. *Plant MolBiol* **102**: 403-416 (2020).
22. Hallgren J, Tsirigos K, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A and Winther O. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv (2022).
23. Chen J, Huang Z, Huang H, Wei S, Liu Y, Jiang C, Zhang J and Zhang C, Selection of relatively exact reference genes for gene expression studies in goosegrass (Eleusine indica) under herbicide stress. *Sci Rep* **7**: 46494 (2017).
24. Livak KJ and Schmittgen TD, Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**: 402-408 (2001).
25. Ritz C, Baty F, Streibig JC and Gerhard D, Dose-response analysis using R. *PLoS One* **10**: 13 (2015).
26. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/. Vienna, Austria (2023).
27. Istiandari P, Yasumoto S, Srisawat P, Tamura K, Chikugo A, Suzuki H, Seki H, Fukushima EO and Muranaka T, Comparative analysis of NADPH-cytochrome P450 reductases from legumes for heterologous production of triterpenoids in transgenic *Saccharomyces cerevisiae*. *Front Plant Sci* **12**: 762546 (2021).
28. Wu D, Shen E, Jiang B, Feng Y, Tang W, Lao S, Jia L, Lin HY, Xie L, Weng X, Dong C, Qian Q, Lin F, Xu H, Lu H, Cutti L, Chen H, Deng S, Guo L, Chuah TS, Song BK, Scarabel L, Qiu J, Zhu QH, Yu Q, Timko MP, Yamaguchi H, Merotto A, Jr., Qiu Y, Olsen KM, Fan L and Ye CY, Genomic insights into the evolution of Echinochloa species as weed and orphan crop. *Nat Commun* **13**: 689 (2022).
29. Aoki D and Yamaguchi H, Genetic relationship between *Echinochloa crus-galli* and *Echinochloa oryzicola* accessions inferred from internal transcribed spacer and chloroplast DNA sequences. *Weed Biol Manage* **8**: 233-242 (2008).
30. Mizutani M and Ohta D, Two isoforms of NADPH: Cytochrome P450 reductase in *Arabidopsis thaliana*: gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol* **116**: 357-367 (1998).
31. Backes WL and Kelley RW, Organization of multiple cytochrome P450s with NADPH-cytochrome P450 reductase in membranes. *Pharmacology & Therapeutics* **98**: 221-233 (2003).
32. Esteves F, Campelo D, Gomes BC, Urban P, Bozonnet S, Lautier T, Rueff J, Truan G and Kranendonk M, The role of the FMN-domain of human cytochrome P450 oxidoreductase in its promiscuous interactions with structurally diverse redox partners. *Frontiers in Pharmacology* **11**2020).

33. Zhu F, Sams S, Moural T, Haynes KF, Potter MF and Palli SR, RNA interference of NADPH-cytochrome P450 reductase results in reduced insecticide resistance in the Bed bug, *Cimex lectularius*. *PLoS One* **7**: e31037 (2012).
34. Adesanya AW, Cardenas A, Lavine MD, Walsh DB, Lavine LC and Zhu F, RNA interference of NADPH-cytochrome P450 reductase increases susceptibilities to multiple acaricides in *Tetranychus urticae*. *Pest Biochem Physiol* **165**: 104550 (2020).
35. Ehlting J, Mattheus N, Aeschliman DS, Li E, Hamberger B, Cullis IF, Zhuang J, Kaneda M, Mansfield SD, Samuels L, Ritland K, Ellis BE, Bohlmann J and Douglas CJ, Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *Plant J* **42**: 618-640 (2005).
36. Sundin L, Vanholme R, Geerinck J, Goeminne G, Höfer R, Kim H, Ralph J and Boerjan W, Mutation of the unducible *Arabidopsis thaliana* cytochrome P450 alters lignin composition and improves saccharification. *Plant Physiol* **166**: 1956-1971 (2014).
37. Niu G, Zhao S, Wang L, Dong W, Liu L and He Y, Structure of the *Arabidopsis thaliana* NADPH-cytochrome P450 reductase 2 (*ATR2*) provides insight into its function. *The FEBS Journal* **284**: 754-765 (2017).

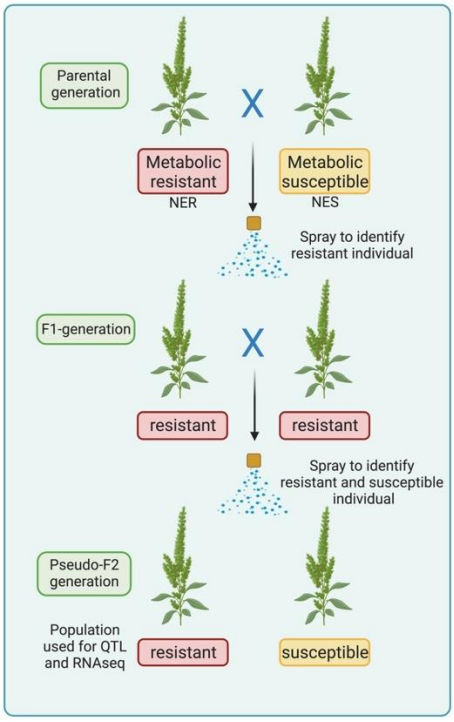
5. APPENDIX A

Supporting Information for CHAPTER II - CYTOCHROME *CYP72A219* IS INVOLVED IN METABOLIC RESISTANCE TO TEMBOTRIONE IN PALMER AMARANTH

Material and Methods

Plant material for QTL mapping and RNAseq experiment

For the crosses, Palmer amaranth seed was sown on 0.7% agar medium (Sigma-Aldrich), placed in a refrigerator at 4 °C for seven d and then germinated on a germination bench at room temperature with 16/8h day/night cycle. Germinated seedlings were transplanted into commercial potting soil (Professional Growing Mix, Sun Gro Horticulture) in 5x5 cm inserts and maintained in the greenhouse at 24±2°C temperatures and 15/9 h day/night photoperiods supplemented with metal-halide lamps (400 µmol m^{-2} s^{-1}). Plants were watered daily. A pseudo-F2 generation was generated by first spraying parental NER individuals at 7-10 cm height with a field rate of 91 a.i. ha^{-1} tembotrione (Laudis, Bayer CropScience) and 1% v/v methylated seed oil (MSO). Herbicide applications were made using an overhead track sprayer (DeVries Manufacturing) equipped with a flat-fan nozzle tip (TeeJet 8002EVS, Spraying System) calibrated to deliver 187 L ha-1 of spray solution at 172 kPa. Surviving NER were transplanted into 22.5 cm diameter pots and individually crossed with another NES individual using pollination bags. Five crosses with NERmale × NESfemale and five crosses with NERfemale × NESmale were performed and grown to seed. The resulting F1 generations were grown out and sprayed again under the same conditions described above. Two F1 individuals from each parental cross that survived the application were crossed with each other. Seeds from the resulting Pseudo-F2 generation were subsequently used for the QTL mapping or RNA-Seq experiment. Schemes of the crosses are available in Supporting Information Figure 3-1.



Supporting Information Figure 3-1. Scheme of crosses performed to generate Pseudo-F2 generation from Palmer amaranth population HPPD-resistant (NER) x susceptible (NES).

RNA-Sequencing

For the RNA-Seq experiment, plants were grown from the Pseudo-F2 generations generated from two parental NERmale x NESfemale crosses (cross A and B, respectively). Seeds were sown on 0.7% agar medium (Sigma-Aldrich), placed in a refrigerator at 4 °C for seven d and then germinated on a germination bench at room temperature with 16/8h of day/night cycle. About 150 seedlings from each Pseudo-F2 cross were transplanted into 4x4 cm inserts each and maintained in a growth chamber at 25/22°C day/night temperatures, 70% relative humidity, and 16h photoperiod with 700 µmol m-2 s-1 provided by incandescent and fluorescent bulbs. At 4-5 cm height and 4-5 leaf stage, the first fully expanded leaf from the apical meristem of each plant was cut and immediately frozen at -80°C for timepoint 0. The plants were then sprayed at 77 g a.i. ha^{-1} (85% of the field rate) with 1% v/v MSO as described above to ensure resistant plants would be able to survive despite removing leaves for testing. Six HAT the second expanded leaf

and twelve HAT the third expanded leaf were cut and immediately frozen at -80 °C, respectively. Visual damage and survival data were recorded after treatment for 21 d to phenotype resistant and susceptible individuals. From each pseudo-F2 cross only the six visually most resistant and the six visually most susceptible individuals were used for RNA-Sequencing.

Total RNA was extracted from frozen ground tissue using the Direct-zol™ RNA MiniPrep Plus (Zymo Research) which includes DNAse treatment. Yield and purity were measured with a NanoDrop 2000 spectrophotometer (Thermo Scientific) and RNA integrity (RIN) was measured on an Agilent 2200 Bio TapeStation system (Agilent Technologies) using Agilent High Sensitivity RNA ScreenTape. RNASeq library preparation was performed with the TruSeq stranded mRNA library prep kit (Illumina) preparing for 150 nucleotide paired-end sequencing. The 108 libraries were run on an Illumina HiSeq 4000 platform on a total of 16 lanes (2 flow cells) with seven libraries per lane, yielding 5.2 billion paired-end reads. Individual library yields were 48.2 million on average and ranged from 35.0 to 71.5 million paired-end reads. On average over 93% of sequenced nucleotides met a quality score of 30 (Q30 Phred score).

Quantitative reverse transcription-PCR (RT-qPCR) validation

Supporting Information Table 3-1. Gene stability assessed by NormFinder algorithm.

Gene name		Stability value
18S		0.110
ACT		0.193
TUB		0.231
Intergroup variation Group identifier	NES	NER
18S	0.068	-0.068
ACT	-0.149	0.149
TUB	0.080	-0.080
Intragroup variation Group identifier	NES	NER
18S	0.515	0.128
ACT	1.048	0.746
TUB	1.147	1.427
Best gene		18S
Stability value		0.110
Best combination of two genes		18S and ACT
Stability value for best combination of two genes		0.112

NES – Nebraska susceptible, NER – Nebraska HPPD-resistant.

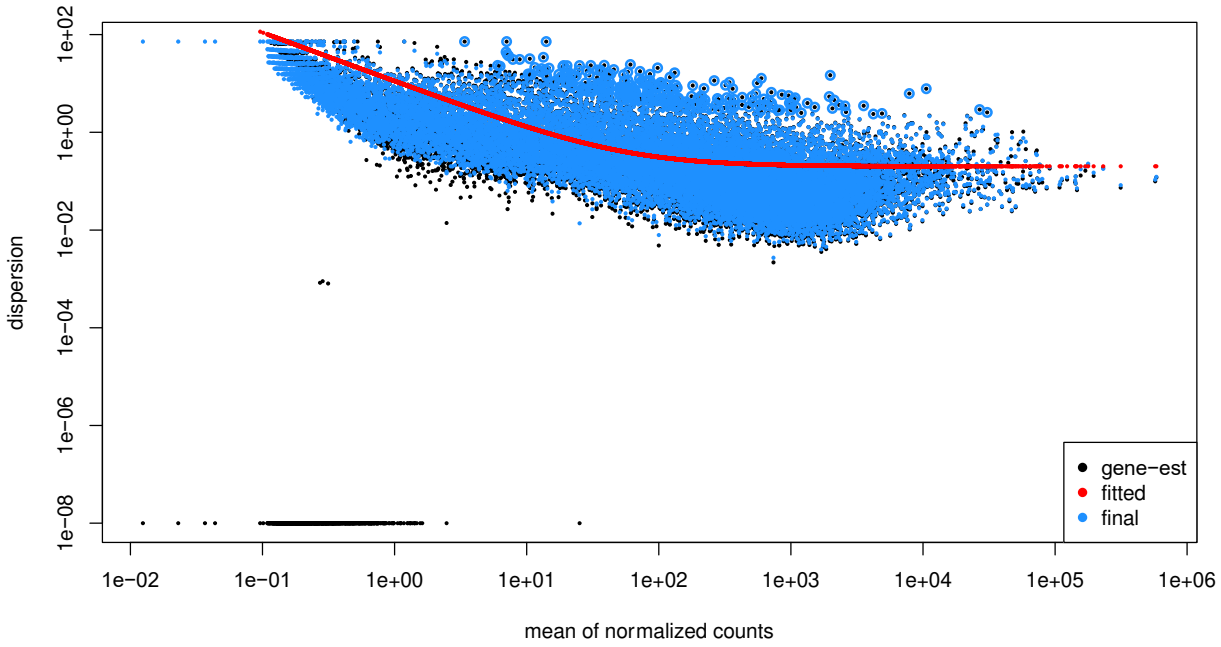
Supporting Information Table 3-2. Primer sequences used for different experiments and their characteristic.

Experiment	gene	sequence 5' to 3'	size (bp)
Gene validation	qPCR_actin_F	ggctgatgcagaggagattc	149
	qPCR_actin_R	gtcccataccaaccatgacacc	
	18S rRNA_F	caaccataaacgatgccgacc	113
	18S rRNA_R	cagccttgcgaccatactcc	
	TUB_F	cattatactgaaggtgctgaac	211
	TUB_R	ctttcggagatgggaaaaca	
	ALSF2 [1]	gctgctgaaggctacgct	118
	ALSR2 [1]	gcg ggactgagtcaagaagtg	
	CYP72A219_4284_F	gccaaaaaggttagaaaaatgc	130
	CYP72A219_4284_R	taggttgggcattgggttta	
	CYP72A219_4285_F	aaagcatcaaaattggcaac	213
	CYP72A219_4285_R	actacgacacctgctggaag	
	CYP72A219_4286_F	tcgaaagccaaaattgaacc	158
	CYP72A219_4286_R	ccttaacactggtgccgaat	
	CYP72A219_7285_F	gccggtgtgcaagttaaaat	232
	CYP72A219_7285_R	atgaaaagtgcggcaaaatc	
	CYP81E8_1F	gacaatgatcgcaccaacac	131
	CYP81E8_1R	cagggaacaatgtggctctt	
Promoter amplification	CYP81E8_F	ggacaaaggagtgaaatcatcg	2131
	CYP81E8_R	catgcggaggaaactaccac	
	CYP72A219_Pro_F	gggtcactgtcttttgatttaggg	1951
	CYP72A219_Pro_R	tggtcatccaagctctttca	
	Promoter_Sen_R	gacattcaggatcctatatattttgttttattgag	-
	Promoter_Sen_250_F_4	gacattcacctgcaggcaattatcaacatcattttctt	250
	Promoter_Sen_750_F_4	gacattcacctgcaggtttttctcaataagattacatt	750
	Promoter_Sen_1500_F_4	gacattcacctgcaggatttaaagcctaatttacattg	1500
	Promoter_Res_R	gacattcaggatcctatatattttatttattgaga	-
	Promoter_Res_250_F_4	gacattcacctgcaggttataatatttagacatgtataat	250
	Promoter_Res_750_F_4	gacattcacctgcaggtttttctcaataattctcaataagtt	750
	Promoter_Res_1500_F_4	gacattcacctgcaggcataatattcttcaattatcaa	1500
	Promoter_chunk_R	gacattcaggatccaaaattgcttggccactcttta	500
	Promoter_chunk_F4	gacattcacctgcagggttttgcagatggaaaagcag	
GUS expression	*NbGAPDH_F* [2]	agctcaagggaattctcgatg	125
	NbGAPDH_R [2]	aaccttaaccatgtcatctccc	
	Nb18S_F [2]	gcaagaccgaaactcaaagg	107
	Nb18S_R [2]	tgttcatatgtcaagggctgg	
	qPCR_GUS_F	ctgatagcgcgtgacaaaaa	172
	qPCR_GUS_R	gagcgtcgcagaacattaca	

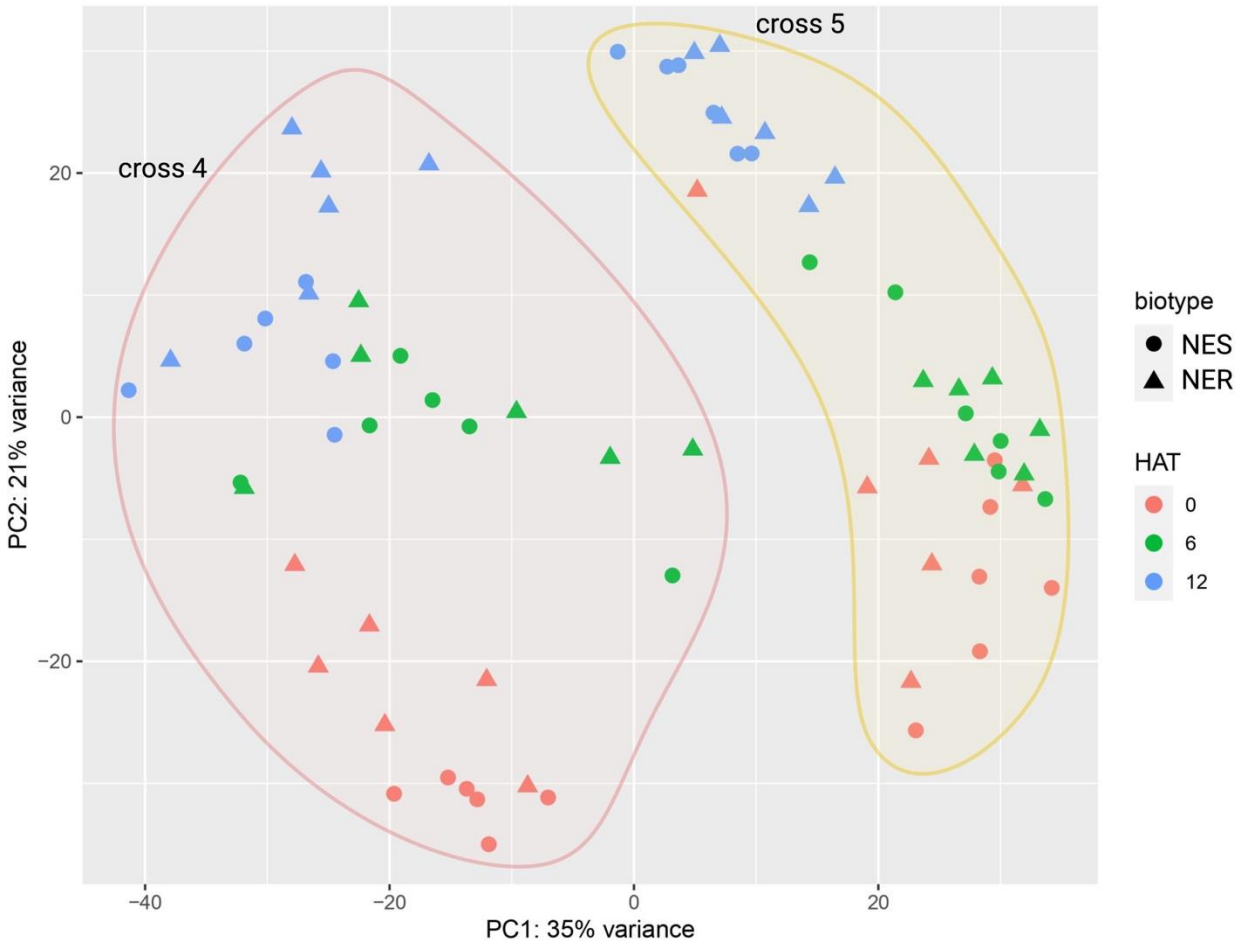
GACATTCA, leader sequence. CCTGCAGG, SbfI restriction site. GGATCC, BamHI restriction site

Results

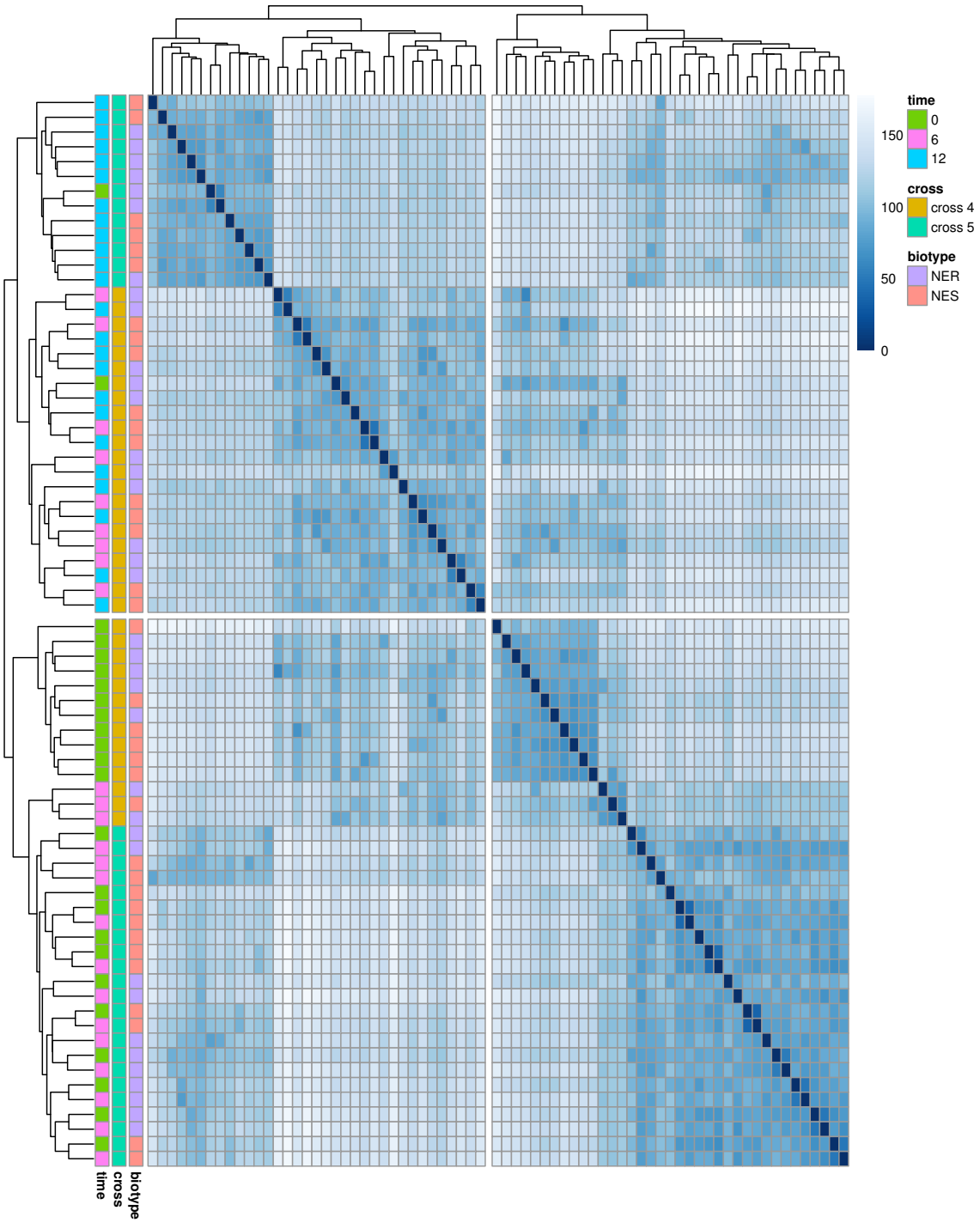
RNA-se analysis



Supporting Information Figure 3-2. Gene-wise dispersion estimates. The dots indicate every gene expression and its estimates toward the fitted curve. Genes with extremely high dispersion are not shrunk toward the curve.



Supporting Information Figure 3-3. Scatterplot of principal components 1 and 2 (35% and 21 % variance explained, respectively) for two different crosses (A and B) between Sensitive (NES) and HPPD-Resistant (NER) *Amaranthus palmeri* population in response to tembotrione before, 6 and 12 HAT.



Supporting Information Figure 3-4. Distance hierarchical clustering heatmap of 72 transcriptomes of Palmer amaranth generated by two crosses (4 and 5) between sensitive (NES) and HPPD-resistant (NER) in response to tembotrione application before, 6 and 12 HAT.

Supporting Information Table 3-3**.** Constitutively differentially expressed genes between tembotrione resistant and susceptible *Amaranthus palmeri*. Adjusted *P*-value <0.01.

gene ID	log2FC	lfcSE	Gene	Molecular function
MAKER_16168	3.57	0.43	Glutathione S-transferase U19 (GSTU9)	glutathione transferase activity
MAKER_12558	5.83	0.92	Disease resistance protein At4g27190-like (LOC104598357)	ADP binding
MAKER_29886	4.20	0.57	NADPH--cytochrome P450 reductase (CYP72A219-like)	monooxygenase activity
MAKER_05008	3.37	0.43	Scopoletin glucosyltransferase (TOGT1)	scopoletin glucosyltransferase activity
MAKER_08144	3.19	0.40	glycosyltransferase	-
MAKER_06269	22.60	3.43	Protein of unknown function	-
MAKER_25717	3.56	0.52	NADPH--cytochrome P450 reductase (CYP72A219-like)	monooxygenase activity
MAKER_08142	3.71	0.56	Protein of unknown function	-
MAKER_21351	2.47	0.33	Crocetin glucosyltransferase (GLT2)	glucosyltransferase activity
MAKER_20679	1.90	0.25	Detoxification 27 (DTX27)	transmembrane transporter activity
MAKER_03299	4.50	0.85	MADS-box transcription factor 23-like	DNA-binding transcription factor activity
MAKER_31185	21.49	4.09	Protein of unknown function	-
MAKER_16324	2.94	0.48	Protein of unknown function	-
MAKER_04610	2.20	0.33	Enhanced pseudomonas susceptibility 1 (EPS1)	acyltransferase activity
MAKER_28341	3.02	0.51	Vacuolar-processing enzyme beta-isozyme (bVPE)	cysteine-type endopeptidase activity
MAKER_20648	2.50	0.40	Betanidin 6-O-glucosyltransferase	glucosyltransferase activity
MAKER_10107	2.64	0.44	Cytochrome P450 81E8 (CYP81E8)	monooxygenase activity
MAKER_07960	2.38	0.39	Glycosyltransferase GTB type superfamily protein	transferase activity
MAKER_08145	2.69	0.45	UDP-glycosyltransferase 71B2-like	UDP-glycosyltransferase activirty
MAKER_04612	2.03	0.31	acetyl-transferase-like	transferase activity
MAKER_32649	4.15	0.89	MADS-box transcription factor 27 (MADS27)	DNA-binding transcription factor activity
MAKER_20644	2.93	0.54	UDP-glycosyltransferase 71A15	transferase activity
MAKER_21756	3.01	0.57	glutathione S-transferase (GST)	transferase activity
MAKER_04611	2.00	0.34	BAHD acyltransferase DCR-like protein	transferase activity
MAKER_15874	3.54	0.80	Protein of unknown function	-
MAKER_06150	2.06	0.36	Glutathione S-transferase U19 (GSTU9)	glutathione transferase activity
MAKER_25718	2.43	0.47	NADPH--cytochrome P450 reductase (CYP72A219-like)	monooxygenase activity
MAKER_24005	2.16	0.42	NADPH:quinone oxidoreductase	NADPH dehydrogenase activity
MAKER_16167	2.16	0.44	Glutathione S-transferase (GST)	glutathione transferase activity
MAKER_34451	3.15	0.95	Protein of unknown function	-
MAKER_17494	3.20	1.05	Protein of unknown function	-
MAKER_12474	3.27	1.90	RING-H2 finger protein ATL3	metal ion binding

MAKER_12523	1.55	0.29	Aldehyde dehydrogenase family 7 member B4 (ALDH7B4)	catalytic activity
MAKER_12487	3.29	1.53	Protein of unknown function	
MAKER_12408	1.81	0.37	Beta-amyrin synthase 1	intramolecular transferase activity
MAKER_22184	2.00	0.46	Calcium-binding protein (CP1)	binding
MAKER_27594	2.61	0.78	Lateral Organ Boundaries (LOB)	DNA binding
MAKER_11763	-2.55	0.4656	Abscisic-aldehyde oxidase (AAO3)	aldehyde oxireductase
MAKER_11325	-1.49	5.1123	Protein of unknown function	-
MAKER_29818	-3.04	1.335	Abscisic acid receptor (PYL4)	abscisic acid binding

FC – fold change, lfcSE – log2fold change standard error for six biological replicates.

CYPs promoter analysis

Alignment of *CYP81E8* promoter of S and R plant

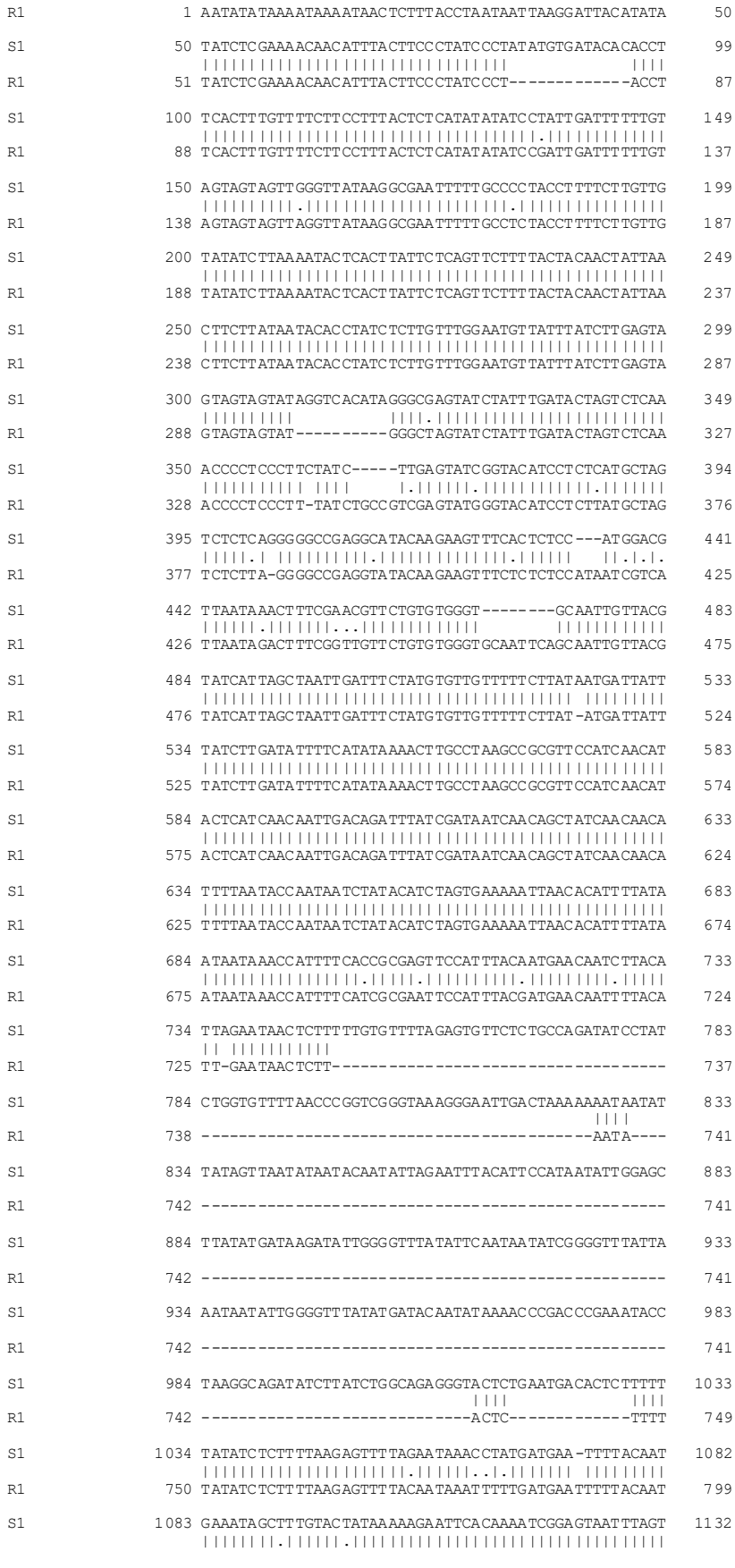
```
#=======================================
#
# Aligned_sequences: 2
# 1: S1
# 2: R1
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1250
# Identity:   1250/1250 (100.0%)
# Similarity:  1250/1250 (100.0%)
# Gaps:        0/1250 ( 0.0%)
# Score: 6250.0
#
#
#=======================================

S1                 1 AAAATACTTCTATAAACTCTTAACACTACCATAAACCAAACCCTACTTTA     50
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
R1                 1 AAAATACTTCTATAAACTCTTAACACTACCATAAACCAAACCCTACTTTA     50

S1                51 CCTACATACCACTCACATAAATATTACGATCTATTTTATTAAGTAAAGAA    100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1                51 CCTACATACCACTCACATAAATATTACGATCTATTTTATTAAGTAAAGAA    100

S1               101 TTCTTTGTTTGTTCCTAAAGAAAACTATAAATTAAATACCTATACTCTCT    150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               101 TTCTTTGTTTGTTCCTAAAGAAAACTATAAATTAAATACCTATACTCTCT    150

S1               151 AAATAACTTTATCTACCATGTGATGGGTGTACCTACGTACATCGGAAAGA    200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               151 AAATAACTTTATCTACCATGTGATGGGTGTACCTACGTACATCGGAAAGA    200

S1               201 AAATCTTATCTCATTTACCATTCATCTTTCTCAGTTTAACTCAAATTTAT    250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               201 AAATCTTATCTCATTTACCATTCATCTTTCTCAGTTTAACTCAAATTTAT    250

S1               251 ACCTATTTACAGAAGTAGTTTAACGTGAATATAACTAATAAGTATCCTTA    300
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               251 ACCTATTTACAGAAGTAGTTTAACGTGAATATAACTAATAAGTATCCTTA    300

S1               301 TACACGTTAATTAAATCTAGACAAATACCCGCCCAATTAGTCTCAGACAG    350
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               301 TACACGTTAATTAAATCTAGACAAATACCCGCCCAATTAGTCTCAGACAG    350

S1               351 GTAAATTAATCCTCTTATCCACTTATTTTTATCTTAAAAATTTCATAATT    400
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               351 GTAAATTAATCCTCTTATCCACTTATTTTTATCTTAAAAATTTCATAATT    400

S1               401 AATCAAAAAACTAACGAATTCTAACACATGAAATCACTCCATAACATCCT    450
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               401 AATCAAAAAACTAACGAATTCTAACACATGAAATCACTCCATAACATCCT    450

S1               451 CCACAATACATACATGAAATTAATATCTCATTAATATCCAAAATTATATC    500
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               451 CCACAATACATACATGAAATTAATATCTCATTAATATCCAAAATTATATC    500
```

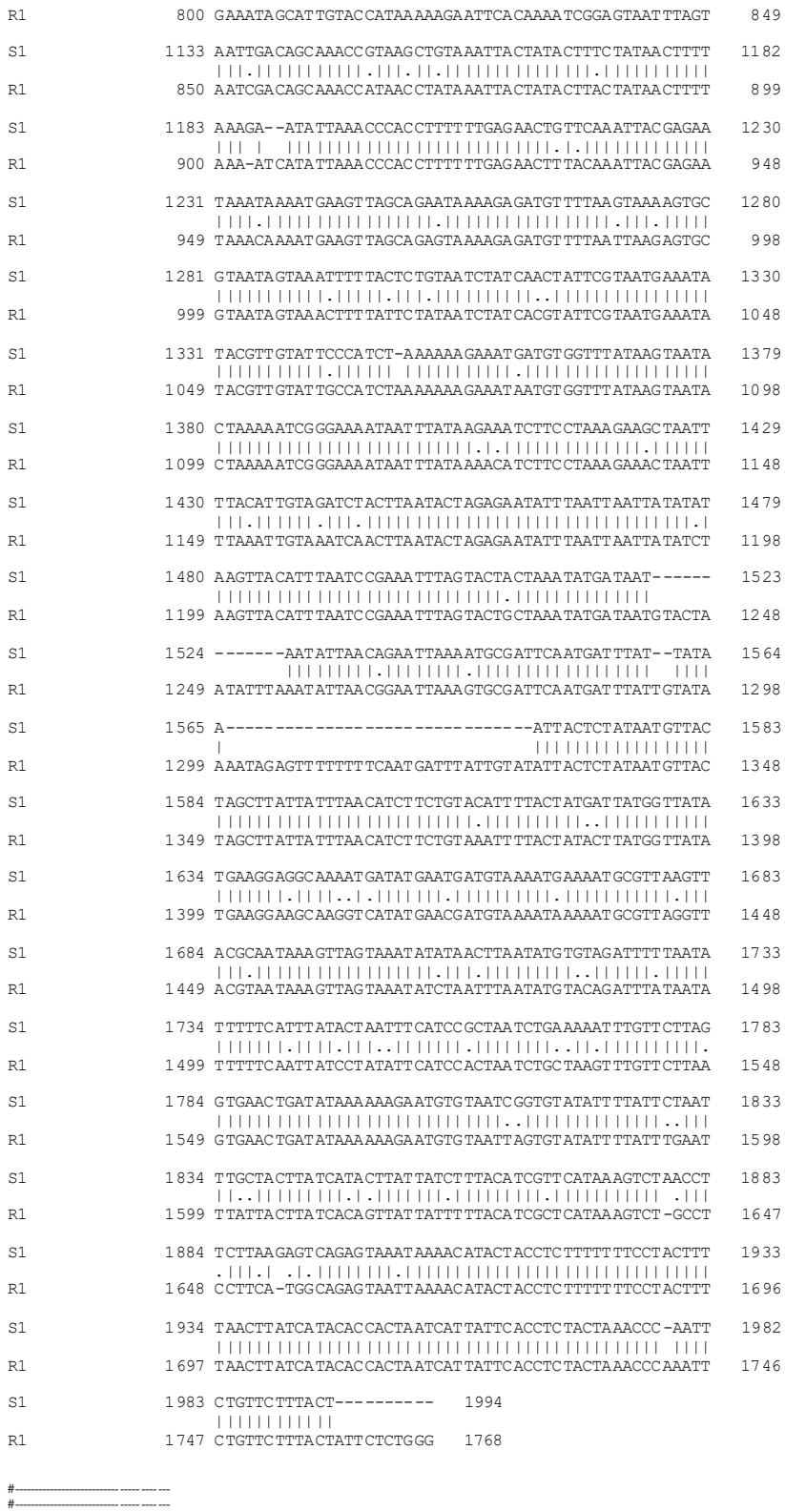
```
S1               501 TTTTCTACTTATTGAAAATCCAAAAACCCTCTATAATATTCATTTAAATA    550
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               501 TTTTCTACTTATTGAAAATCCAAAAACCCTCTATAATATTCATTTAAATA    550

S1               551 CATACTTTCCCATTACCCCGTTTAATATACCTGTTTGATTTTAACATATA    600
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               551 CATACTTTCCCATTACCCCGTTTAATATACCTGTTTGATTTTAACATATA    600

S1               601 TCTCGTTTTAATTTCTCTAGTGAATTAAATCCTCTATAAAAGAGTGCATT    650
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               601 TCTCGTTTTAATTTCTCTAGTGAATTAAATCCTCTATAAAAGAGTGCATT    650

S1               651 TTTAGTACACAATTTTTACACTTACGTTTAGTACAGTAAGAGTATGTGCT    700
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               651 TTTAGTACACAATTTTTACACTTACGTTTAGTACAGTAAGAGTATGTGCT    700

S1               701 AAAGTATAAAAAGTGCATTTTATTAATAATTCCAACAGAAGATTATAGTA    750
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               701 AAAGTATAAAAAGTGCATTTTATTAATAATTCCAACAGAAGATTATAGTA    750

S1               751 TAATTCTTTGATTAATTAATCGGTACTCAAAGTCAAATACCAACATTGAG    800
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               751 TAATTCTTTGATTAATTAATCGGTACTCAAAGTCAAATACCAACATTGAG    800

S1               801 ACTATACAAATATAATAGAGTTTGAATATTTGAGATCAGGTACTATACGA    850
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               801 ACTATACAAATATAATAGAGTTTGAATATTTGAGATCAGGTACTATACGA    850

S1               851 TTATAAGCTAGTTCAAATATATCAAGTTAATATTCCTCAAAAGTATATGA    900
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               851 TTATAAGCTAGTTCAAATATATCAAGTTAATATTCCTCAAAAGTATATGA    900

S1               901 CATAACGAATTAAATCGATGAGAATACTCTCTAGCAGAAAATTTCTATGT    950
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               901 CATAACGAATTAAATCGATGAGAATACTCTCTAGCAGAAAATTTCTATGT    950

S1               951 TAAGTATTTATTATTCAGGTGTAAGAATAAAAGAGAAATTAAACATAATT   1000
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1               951 TAAGTATTTATTATTCAGGTGTAAGAATAAAAGAGAAATTAAACATAATT   1000

S1              1001 CACCCGGTAAATCGGGTATTTATATTATAACACAGAGAGCCTCTCTGGCA   1050
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1              1001 CACCCGGTAAATCGGGTATTTATATTATAACACAGAGAGCCTCTCTGGCA   1050

S1              1051 GCGAGTGTTGTTAAACACCATAACGATATAATAGTTGTTTCTGTAATTAT   1100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1              1051 GCGAGTGTTGTTAAACACCATAACGATATAATAGTTGTTTCTGTAATTAT   1100

S1              1101 AACTATATCATTATGTTGTCTTAATATTTTTCTCTTGTTGAATTATAAAA   1150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1              1101 AACTATATCATTATGTTGTCTTAATATTTTTCTCTTGTTGAATTATAAAA   1150

S1              1151 TACTTTTTTATACTCTTTATATTTTTTCATTTTTTTTAGTTTATATCAAT   1200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1              1151 TACTTTTTTATACTCTTTATATTTTTTCATTTTTTTTAGTTTATATCAAT   1200

S1              1201 TTTATATTCATCGTTTAAATTATCTTGTCACTTTTTTTAACTTTATTTCA   1250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
R1              1201 TTTATATTCATCGTTTAAATTATCTTGTCACTTTTTTTAACTTTATTTCA   1250
```

#---------------------------------------
#---------------------------------------

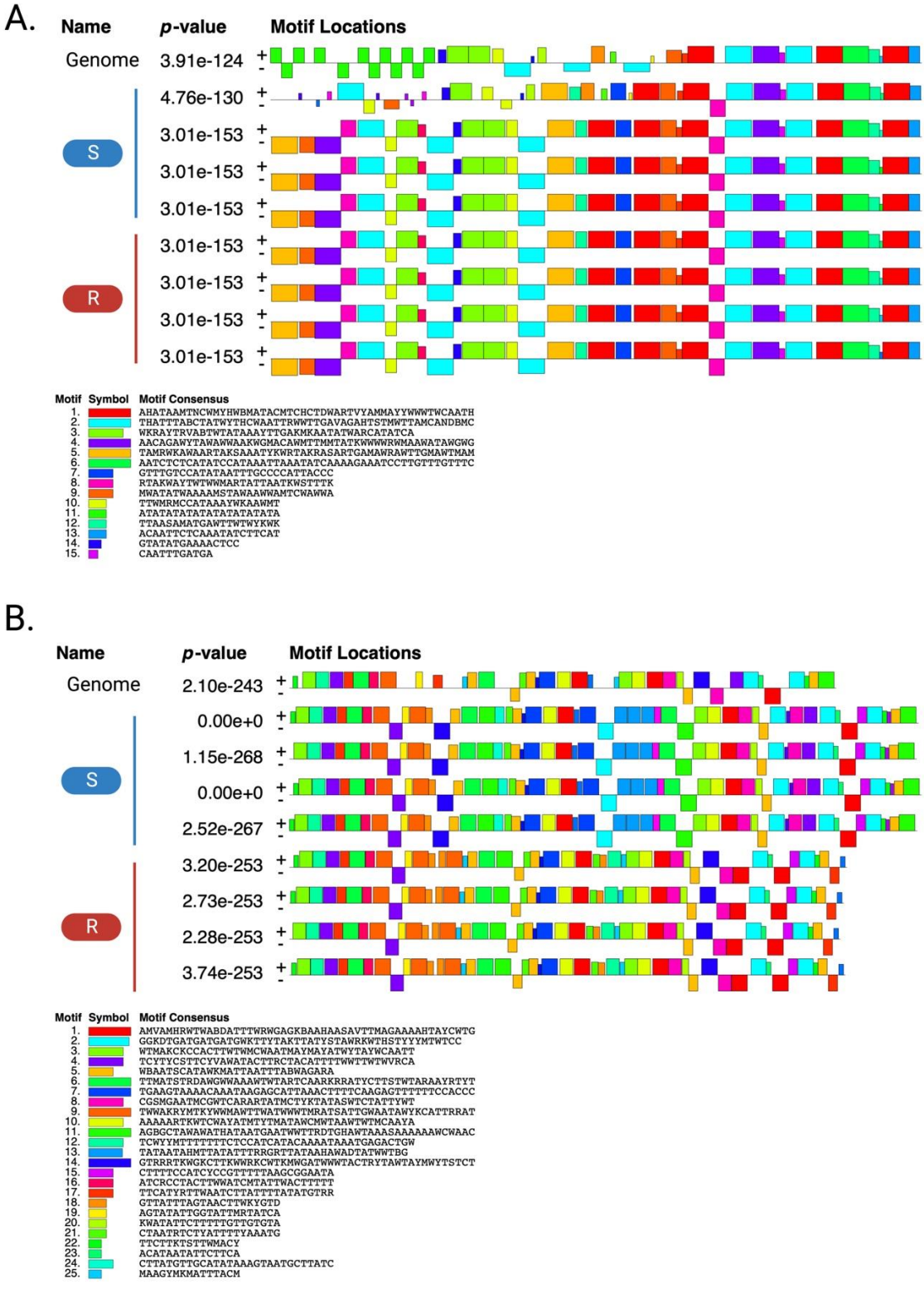
Alignment of *CYP72A219_4284* promoter of S and R plant

```
#=======================================
#
# Aligned_sequences: 2
# 1: S1
# 2: R1
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 2072
# Identity:    1593/2072 (76.9%)
# Similarity:  1593/2072 (76.9%)
# Gaps:         382/2072 (18.4%)
# Score: 7172.5
#
#
#=======================================

S1                 1 AATATATAAAACAAAATAACTCTTTACCTAATAATTAAGGATTA-ATATA     49
                     ||||||||||.|||||||||||||||||||||||||||||| |||||
```





Supporting Information Figure 3-5. Needle alignment of *CYP81E8* and *CYP72A219_4284* promoter between S and R sample.



Supporting Information Figure 3-6. Motifs analysis using MEME-suite tool [3] for *CYP81E8* and *CYP72A219_4284* promoter between S and R samples. Different color indicates different motif.

Involvement of *CYP72A219_4284* in different HPPD-resistant Palmer amaranth populations

Supporting Information Table 3-4. Logistic equation parameters and the resistance index (RI) of survival (% of untreated control) for different populations of Palmer amaranth subjected to different doses of tembotrione.

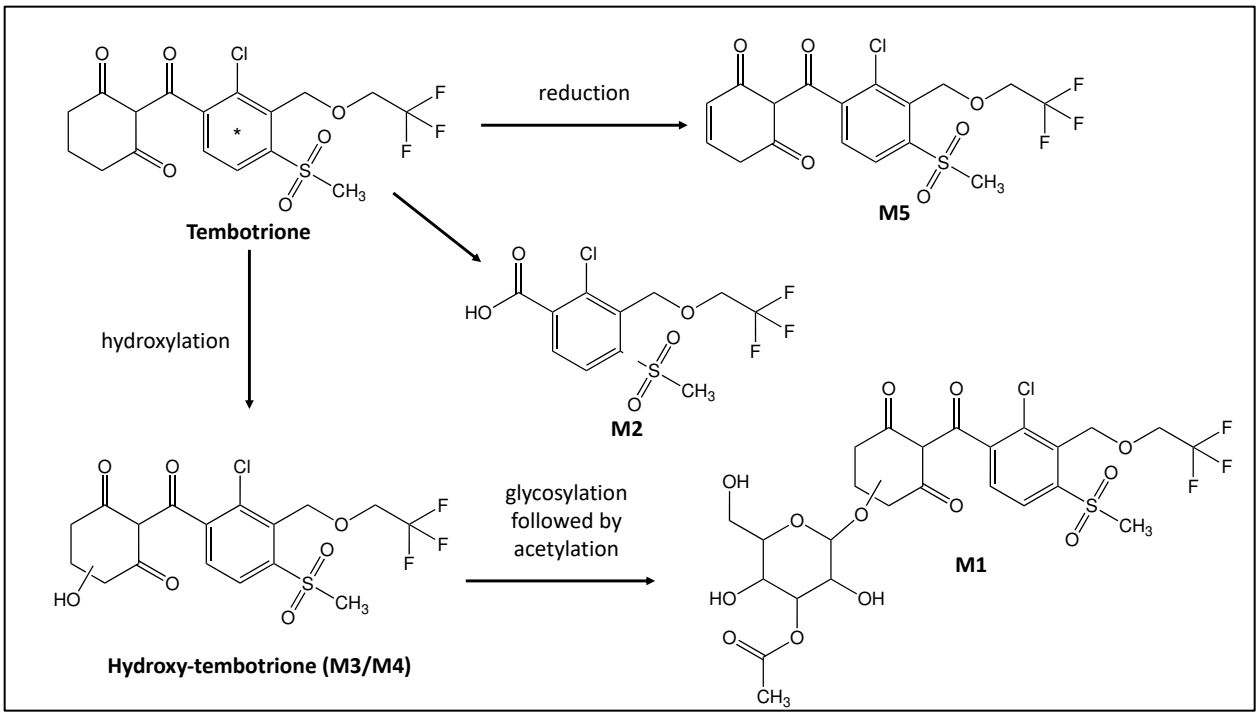
populations	b	d	LD_{50}				RI	*p*-value
			dose (g ha^{-1})	lower CI (95%)	upper CI (95%)	SE		
Sensitive	12.6	100.0	22.8	21.8	23.8	0.5	-	
NER	1.9	98.8	99.9	74.5	125.5	12.9	4.3	<0.001
WR2013-034	2.9	98.5	89.4	73.1	105.8	8.3	3.9	<0.001
W2019-044	2.5	99.5	77.3	61.6	93.0	8.0	3.4	<0.001
W2019-137	11.6	100.0	181.9	172.8	191.1	4.6	7.9	<0.001
W2019-140	3.8	100.3	129.4	108.2	150.6	10.8	5.7	<0.001
W2019-141	4.1	98.9	55.4	46.8	64.5	4.4	2.4	<0.001
W2019-144	5.2	100.0	69.6	57.8	81.3	5.9	3.0	<0.001
W2019-198	6.2	97.3	105.2	87.2	123.24	9.1	4.6	<0.001
W2019-199	2.1	99.5	132.1	100.4	163.2	15.8	5.8	<0.001
W2019-200	1.5	99.0	178.3	127.6	229.6	25.9	7.8	<0.001
W2019-273	2.8	100.6	63.8	51.9	75.9	6.1	2.8	<0.001
W2019-274	2.8	99.5	70.3	57.1	83.6	6.7	3.1	<0.001

b: slope; d: upper limit; LD_{50}: herbicide dose that causes a 50% reduction in survival, CI: confidence interval of the parameter LD_{50} ($\alpha = 0.05$); and RI: resistance index = LD_{50} ratio between respective population and sensitive.

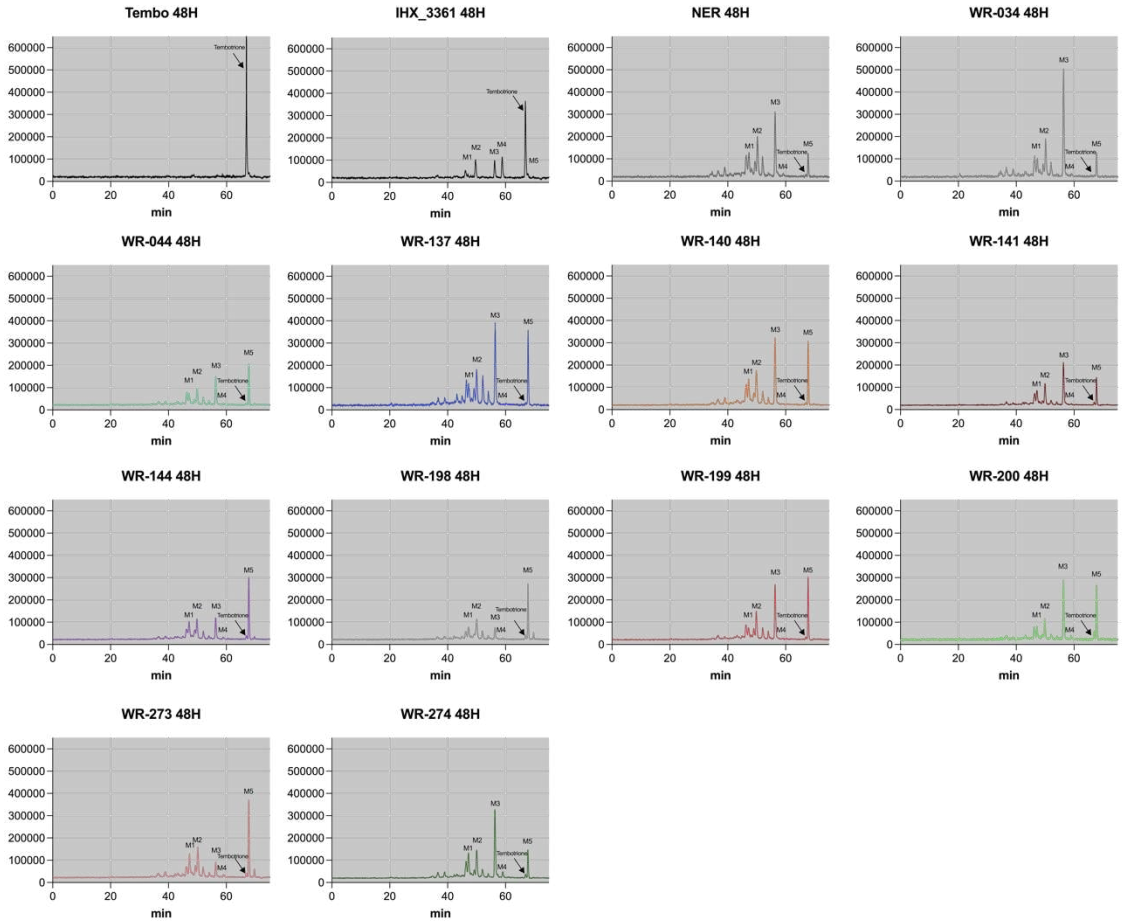
Supporting Information Table 3-5. Logistic equation parameters and the resistance index (RI) of shoot fresh weight (of untreated control) for different populations of Palmer amaranth subjected to different doses of tembotrione.

Populations	b	d	GR50				RI	*p*-value
			dose (g ha^{-1})	lower CI (95%)	upper CI (95%)	SE		
Sensitive	1.5	100	7.3	4.3	10.4	1.5	-	-
NER	2.1	100	50.6	32.2	69.0	9.3	6.9	<0.01
WR2013-034	5.7	100	40.1	29.2	51.1	5.5	5.5	<0.01
W2019-044	17.9	100	39.9	0.65	79.1	19.9	5.4	0.1
W2019-137	1.5	100	50.2	29.6	70.6	10.4	6.8	<0.01
W2019-140	2.4	100	43.2	29.6	70.6	6.9	5.9	<0.01
W2019-141	2.5	100	27.5	18.9	36.1	4.4	3.7	<0.01
W2019-144	4.1	100	20.3	15.6	25.0	2.4	2.8	<0.01
W2019-198	2.6	100	24.9	17.4	32.4	3.8	3.4	<0.01
W2019-199	2.1	100	19.8	9.5	30.0	5.2	2.7	0.06
W2019-200	4.5	100	31.9	23.9	39.8	4.0	4.3	<0.01
W2019-273	5.1	100	24.1	17.8	30.2	3.1	3.3	<0.01
W2019-274	3.3	100	29.2	20.6	37.8	4.4	3.9	<0.01

b: slope; d: upper limit; GR_{50}: herbicide dose that causes a 50% reduction in shoot fresh weight, CI: confidence interval of the parameter GR_{50} ($\alpha = 0.05$); and RI: resistance index = GR_{50} ratio between respective population and sensitive.

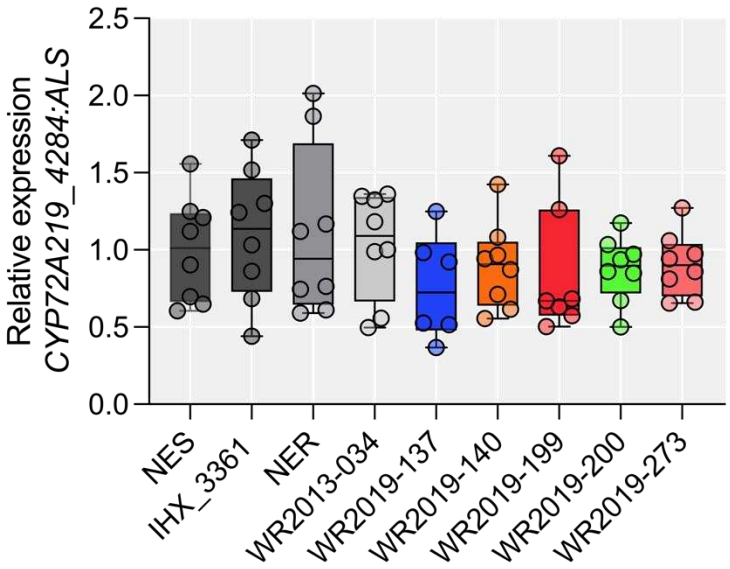


Supplemental Information Figure 3-7**.** Chemical structures of tembotrione and the main metabolites. Detoxification pathways proposed by Küpper et al. (2018). The asterisk in the tembotrione molecule marks the location of the ^{14}C-label.



Supporting Information Figure 3-8. Representative reverse-phase HPLC chromatogram for Palmer populations 48 HAT with ^{14}C-Tembotrione. Retention times of 47.3, 50.1, 56.3, 59.0, 66.9 and 67.7 min correspond to M1, M2, M3, M4, tembotrione and M5, respectively. IHX_3361 is herbicide sensitive control, NER and WR2013-034 are HPPD resistant control.

Gene copy number



Supporting Information Figure 3-9. *CYP72A219_4284* copy number relative to *ALS* in different populations of *Amaranthus palmeri*. NES and IHX_3361 are herbicide sensitive control, NER and WR2013-034 are HPPD resistant control. Circle indicates biological replicate.

Degree of dominance and QTL identification

Supporting Information Table 3-6. Quantitative trait loci found in Pseudo-F2 *Amaranthus palmeri* in cross B and combined cross A + B for the trait of HPPD resistance. Position are based on available genome of Palmer amaranth[4].

QTL	LOD score	Position	Interval		Chromosome	Annotated genes	PVE
Cross B							
Scaffold_10	11.6	11877263	11865155	11890395	4	77	23 %
	8.4	10336852	10224431	10336879			
	7.9	10337656	10336879	10482794			
Scaffold_81	10.1	19632215	18232833	20375576	2	78	20 %
	8.6	18232833	18186955	18238951			
Scaffold_6	9.5	16851238	15842545	16851238	8	7	19 %
	8.7	15843735	15752153	15907859			
Scaffold_14	8.0	7800824	2267056	7800824	15	50	16 %
	7.9	4777664	4754252	4917017			
	7.9	2267056	2226097	2271946			
combined cross A + B							
Scaffold_10	12.25	11877263	10877263	12877263	4	105	15 %
	7.84	11332443					
Scaffold_6	7.81	5462074	4962074	5962074	8	32	10 %
	7.60	16851238	15150809	17351238		164	
	7.00	15650809					
	6.68	15617367					
Scaffold_14	7.87	7800824	7300824	8300824	15	25	10 %

PVE – phenotype variation explained.

Cytochrome P450 sequences and alignments

CYP72A219_4284

NES and NER CYP72A219_4284 alleles from sanger sequencing
Reference genome use for the alignment:
palmeri(v1.1, unmasked), Name: mRNA4451, Type: CDS, Feature Location: (Chr: Scaffold_10, join(13520341..13520623,13521428..13521645,13522474..13522718,13523654..13524035,13525026..13525451)) Genomic Location: 13520341-13525451

>NES_CYP71A219_5284 allele1
ATGGAAGAGACAAGCATATTATCAATTGGAATATCTTTGGTTTGTGTAGTATTAGTAAGTGTCTCATGGGGAATTCTAAATTGGGT
ATGGTTTAAGCCAAAAAGGTTAGAAAAATGCCTTAGAAATCAAGGGTTTAATGGAAATTCTTATAGGTTATTGTATGGAGATACC
AATGAAAGAGCTAGGATGACTATGGAATCTAGATTAAACCCAATGCCTAACCTATCCAATCATTTTCTTCCACATTCTCTTGCATT
TTACCATCACATTATAACTAAGTATGGGAGAAGATGCTTTATATGGAATGGACCTATTCCATTGGTAAACATTGAAGAACCAAATT
TAGTAAGAGAAATATTTATGAAAATAAATGATTTTCAAAAAGCAAAGTCAAACCCAATCATAAAGAAGGTTGCACCGGGCTTGCT
TTTCTTAGAGGGCCATGCATGGGCTCAACGAAGGAAACTCCTCACTCCTGCTTTCCACATCGATAAGCTCAAGTTTATGGTTCCAA
CAATATGGGAAAGTTCAGAAGAAATGATTAAAGAATGGGAAGAATTGGTATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGT
CATCTTTACAAAAATTATCTGCAGATATGATTTCTCGAGCAGCATTTGGAAGTAGTTATCAAGAAGGGAAGACAATTTTTGATCTT
TTAACTGAACAAATTAAGATTGTTGTTCCAGTATTTAATTCTGTTTATATCCCTGGTTGGAGGTTTTTGCCAACAAAAACAAATCG
AAGAATATCAGAAATCGACAGAGAAATCAGGATTCGACTAAAAGGAATAATCGACAAAAGAAAGAAATACGATAAATAGAGAAG
AGAGGGCAAAGGATGACTTGCTTAGTACATTAATTCAAGACAATTTTGGACATGCAAATAAAAAGCAAGACATAAATAAGTTAA
ATATTGAAGAAATTGTTAATGAATGTAAATTATTTTACTTAGCCGGACAAGATACAACTTCTTCATTATTGTTATGGACACTAATT
ATGTTGGGAAAACACCAAATATGGCAACAAAGAGCAAGAGAAGAGATCATCAATACATTTGGGGATGCCAAACCTCAATTTCAT
GAGTTGAATAGTCTTAAAATAGTAAACATGATACTACAAGAGGTTCTAAGATTATATCCACCAGTATTAGAATTAAATCGTGAAG
TTACTCAAGACATAAAAATAGGAGATGTAATACTTCCGGCAGGAGTTTTACTTAATATACCAATTTTACTTTTGCAACAAGATGAA
AAAATATGGGGTAAAGATGCGAAAGAATTTAATCCCGAAAGATTTAGTGAAGGAATTTCAAAAGCTAGTAAAGGAAATATGTCA
TTCTTTGCGTTCGGTTGGGGACCTAGAATATGTCTTGGATCGAATTTTGCTATGATTGAGGCCAAATTAACTTTAGTCTTCATTCTG
CAAAGGTTTGTATTTGAGCTTTCACCATCTTATCGTCATGCCCCTTTGAATGTTGGAACTCTTAGACCTCAATTTGGTGCACCAATT
ATATTTCGTCGACGATAA

>NES_CYP71A219_5284 allele2
ATGGAAGAGACAAGCATATTATCAATTGGAATATCTTTGGTTTGTGTAGTATTATTAAGTGTCTCATGGGTAATTCTAAATTGGGT
ATGGTTTAAGCCAAAAAGGTTAGAAAAATGCCTTAGAAATCAAGGGTTTAATGGAAATTCTTATAGGTTATTGTATGGAGATACC
AATGAAAGAGCTAGGATGACCATGGAATCTAGATTAAACCCAATGCCTAACCTATCCAATCATTTTCTTCCACATTCTCTTGCAT T
TTTCCATCACATTATAACTAAATATGGGAGAAGATGCTTTATATGGAATGGACCTATTCCATTGGTAAACATAGAAGAACCATATT
TAATAAGAGAAATATTTATGAAAATAAATGATTTTCAAAAAGCAAAGTCAAACCCAATCATAAAGAAGGTTGCACCGGGCTTGCT
TTTCTTAGAGGGCCATGCATGGGCTCAACGAAGGAAACTCCTCACTCCTGCTTTCCACATAGATAAGCTCAAGTTTATGGTTCCAA
CAATATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGT
CATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCTGCTTTTGGAAGTAGTTATGAAGAAGGGAAGACAATTTTTGATCTT
TTAACTGAACAAATTAAGATTGTAGTTCCAGTATTTAATTCTGTTTATATCCCTGGTTGGAGGTTTTTACCAACCAAAACAAATCG
AAGAATATCAGAAATCGACAGAGAAATCAGGATCCGACTAAAAGGAATAATCGACAAAAGAAAGAAATACGATAAATAGAGAAG
AGAGGGCAAAGGATGACTTGCTTAGTACATTAATTCAAGACAATTTTGGACATGCAAATAAAAAGCAAGACATAAATAAATTAA
GTATTGAAGAAGTTGTTAATGAATGTAAATTATTTTACTTAGCTGGACAAGATACAACTTCTTCCTTATTGTTATGGACACTAATTA
TGTTGGGAAAGCACCAAATATGGCAACAAAGAGCAAGAGAAGAGATCATCAATGCATTTGGGGATGCCAAACCTCAATTTCATG
AGTTGAATAGTCTTAAAATAGTAAACATGATACTACAAGAGGTTCTAAGATTATATCCACCAATATTAGAATTAAATCGTGAAGT

>NER_CYP71A219_5284 allele1
ATGGAAGAGACAAGCATATTATCAATTGGAATATCTTTGGTTTGTGTAGTATTATTAAGTGTCTCATGGGTAATTCTAAATTGGGT
ATGGTTTAAGCCAAAAAGGTTAGAAAAATGCCTTAGAAATCAAGGGTTTAATGGAAATTCTTATAGGTTATTGTATGGAGATACC
AATGAAAGAGCTAGGATGACCATGGAATCTAGATTAAACCCAATGCCTAACCTATCCAATCATTTTCTTCCACATTCTCTTGCATT
TTTCCATCACATTATAACTAAATATGGGAGAAGATGCTTTATATGGAATGGACCTATTCCATTGGTAAACATAGAAGAACCATATT
TAATAAGAGAAATATTTATGAAAATAAATGATTTTCAAAAAGCAAAGTCAAACCCAATCATAAAGAAGGTTGCACCGGGCTTGCT
TTTCTTAGAGGGCCATGCATGGGCTCAACGAAGGAAACTCCTCACTCCTGCTTTCCACATAGATAAGCTCAAGTTTATGGTTCCAA
CAATATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGT
CATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCTGCTTTTGGAAGTAGTTATGAAGAAGGGAAGACAATTTTTGATCTT
TTAACTGAACAAATTAAGATTGTAGTTCCAGTATTTAATTCTGTTTATATCCCTGGTTGGAGGTTTTTACCAACCAAAACAAATCG
AAGAATATCAGAAATCGACAGAGAAATCAGGATCCGACTAAAAGGAATAATCGACAAAAGAAAGAAATACGATAAATAGAGAAG
AGAGGGCAAAGGATGACTTGCTTAGTACATTAATTCAAGACAATTTTGGACATGCAAATAAAAAGCAAGACATAAATAAATTAA
GTATTGAAGAAGTTGTTAATGAATGTAAATTATTTTACTTAGCTGGACAAGATACAACTTCTTCCTTATTGTTATGGACACTAATTA
TGTTGGGAAAGCACCAAATATGGCAACAAAGAGCAAGAGAAGAGATCATCAATGCATTTGGGGATGCCAAACCTCAATTTCATG
AGTTGAATAGTCTTAAAATAGTAAACATGATACTACAAGAGGTTCTAAGATTATATCCACCAATATTAGAATTAAATCGTGAAGT

TACTCAAGACATAAAAATAGGAGATGTAATACTTCCGGCAGGAGTTTTACTTAATATACCAATTTTACTTTTGCAACAAGATGAAA
AAATATGGGGTAAAGATGCGAAAGAATTTAATCCCGAAAGATTTAGTGAAGGAATTTCAAAAGCTAGTAAAGGAAATATGTCAT
TCTTTGCATTCGGTTGGGGACCTAGAATATGTCTTGGATCGAATTTTGCTATGATTGAGGCCAAATTAACTTTAGTCTTCATTCTG C
AAAGGTTTGTATTTGAGCTTTCACCATCTTATCGTCATGCCCCTTTGAATGTTGGAACTCTTAGACCTCAATTTGGTGCACCAATTA
TATTTCGTCGACGATAA

>NER_CYP71A219_5284 allele2
ATGGAAGAGACAAGCATATTATCAATTGGAATATCTTTGGTTTGTGTAGTATTAGTAAGTGTCTCATGGGGAATTCTAAATTGGGT
ATGGTTTAAGCCAAAAAGGTTAGAAAAATGCCTTAGAAATCAAGGGTTTAATGGAAATTCTTATAGGTTATTGTATGGAGATACC
AATGAAAGAGCTAGGATGACTATGGAATCTAGATTAAACCCAATGCCTAACCTATCCAATCATTTTCTTCCACATTCTCTTGCATT
TTACCATCACATTATAACTAAGTATGGGAGAAGATGCTTTATATGGAATGGACCTATTCCATTGGTAAACATTGAAGAACCAAATT
TAGTAAGAGAAATATTTATGAAAATAAATGATTTTCAAAAAGCAAAGTCAAACCCAATCATAAAGAAGGTTGCACCGGGCTTGCT
TTTCTTAGAGGGCCATGCATGGGCTCAACGAAGGAAACTCCTCACTCCTGCTTTCCACATCGATAAGCTCAAGTTTATGGTTCCAA
CAATATGGGAAAGTTCAGAAGAAATGATTAAAGAATGGGAAGAATTGGTATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGT
CATCTTTACAAAAATTATCTGCAGATATGATTTCTCGAGCAGCATTTGGAAGTAGTTATCAAGAAGGGAAGACAATTTTTGATCTT
TTAACTGAACAAATTAAGATTGTTGTTCCAGTATTTAATTCTGTTTATATCCCTGGTTGGAGGTTTTTGCCAACAAAAACAAATCG
AAGAATATCAGAAATCGACAGAGAAATCAGGATTCGACTAAAAGSAATAATCGACAAAAGAAAGAATACGATAAATAGAGAAG
AGAGGGCAAAGGATGACTTGCTTAGTACATTAATTCAAGACAATTTTGGACATGCAAATAAAAAGCAAGACATAAATAAGTTAA
ATATTGAAGAAATTGTTAATGAATGTAAATTATTTTACTTAGCCGGACAAGATACAACTTCTTCATTATTGTTATGGACACTAATT
ATGTTGGGAAAACACCAAATATGGCAACAAAGAGCAAGAGAAGAGATCATCAATACATTTGGGGATGCCAAACCTCAATTTCAT
GAGTTGAATAGTCTTAAAATAGTAAACATGATACTACAAGAGGTTCTAAGATTATATCCACCAGTATTAGAATTAAATCGTGAAG
TTACTCAAGACATAAAAATAGGAGATGTAATACTTCCGGCAGGAGTTTTACTTAATATACCAATTTTACTTTTGCAACAAGATGAA
AAAATATGGGGTAAAGATGCGAAAGAATTTAATCCCGAAAGATTTAGTGAAGGAATTTCAAAAGCTAGTAAAGGAAATATGTCA
TTCTTTGCGTTCGGTTGGGGACCTAGAATATGTCTTGGATCGAATTTTGCTATGATTGAGGCCAAATTAACTTTAGTCTTCATTCTG
CAAAGGTTTGTATTTGAGCTTTCACCATCTTATCGTCATGCCCCTTTGAATGTTGGAACTCTTAGACCTCAATTTGGTGCACCAATT
ATATTTCGTCGACGATAA

CLUSTAL alignment

```
CYP72A219_4284 genom MEETSILSIGISLVCVVLVSVSWGILNWVWFKPKRLEKCLKNQGFNGNSYRLLYGDTNER   60
NES allele1          MEETSILSIGISLVCVVLLSVSWVILNWVWFKPKRLEKCLRNQGFNGNSYRLLYGDTNER   60
NES allele2          MEETSILSIGISLVCVVLVSVSWGILNWVWFKPKRLEKCLRNQGFNGNSYRLLYGDTNER   60
NER allele1          MEETSILSIGISLVCVVLLSVSWVILNWVWFKPKRLEKCLRNQGFNGNSYRLLYGDTNER   60
NER allele2          MEETSILSIGISLVCVVLVSVSWGILNWVWFKPKRLEKCLRNQGFNGNSYRLLYGDTNER   60

CYP72A219_4284 genom AWMTMESRLNPMPNLSNHFLPHSLAFFHHIITKYGRRCFIWNGPIPLVNIEEPNLIREIF  120
NES allele1          ARMTMESRLNPMPNLSNHFLPHSLAFFHHIITKYGRRCFIWNGPIPLVNIEEPYLIREIF  120
NES allele2          ARMTMESRLNPMPNLSNHFLPHSLAFYHHIITKYGRRCFIWNGPIPLVNIEEPNLVREIF  120
NER allele1          ARMTMESRLNPMPNLSNHFLPHSLAFFHHIITKYGRRCFIWNGPIPLVNIEEPYLIREIF  120
NER allele2          ARMTMESRLNPMPNLSNHFLPHSLAFYHHIITKYGRRCFIWNGPIPLVNIEEPNLVREIF  120

CYP72A219_4284 genom MKINDFQKVKSNPIIKKVAPGLLFLEGHAWAQRRKLLTPAFHIDKLKFMVPTIWESSNEM  180
NES allele1          MKINDFQKAKSNPIIKKVAPGLLFLEGHAWAQRRKLLTPAFHIDKLKFMVPTIWESSKEM  180
NES allele2          MKINDFQKAKSNPIIKKVAPGLLFLEGHAWAQRRKLLTPAFHIDKLKFMVPTIWESSEEM  180
NER allele1          MKINDFQKAKSNPIIKKVAPGLLFLEGHAWAQRRKLLTPAFHIDKLKFMVPTIWESSKEM  180
NER allele2          MKINDFQKAKSNPIIKKVAPGLLFLEGHAWAQRRKLLTPAFHIDKLKFMVPTIWESSEEM  180

CYP72A219_4284 genom IKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYEEGKTIFDLLTEQIKIVVPVF  240
NES allele1          IKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYEEGKTIFDLLTEQIKIVVPVF  240
NES allele2          IKEWEELVSKTGSCEIEVWSSLQKLSADMISRAAFGSSYQEGKTIFDLLTEQIKIVVPVF  240
NER allele1          IKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYEEGKTIFDLLTEQIKIVVPVF  240
NER allele2          IKEWEELVSKTGSCEIEVWSSLQKLSADMISRAAFGSSYQEGKTIFDLLTEQIKIVVPVF  240

CYP72A219_4284 genom NSVYIPGWRFLPTKTNRRISEIDREIRIRLKGIIDKRKNTINREERAKDDLLSTLIQDNF  300
NES allele1          NSVYIPGWRFLPTKTNRRISEIDREIRIRLKGIIDKRKNTINREERAKDDLLSTLIQDNF  300
NES allele2          NSVYIPGWRFLPTKTNRRISEIDREIRIRLKGIIDKRKNTINREERAKDDLLSTLIQDNF  300
NER allele1          NSVYIPGWRFLPTKTNRRISEIDREIRIRLKGIIDKRKNTINREERAKDDLLSTLIQDNF  300
NER allele2          NSVYIPGWRFLPTKTNRRISEIDREIRIRLKGIIDKRKNTINREERAKDDLLSTLIQDNF  300

CYP72A219_4284 genom GHANKKQDINKLNIEEVVNECKLFYLAGQDTTSSLLLWTLIMLGKHQIWQQRAREEIINT  360
NES allele1          GHANKKQDINKLSIEEVVNECKLFYLAGQDTTSSLLLWTLIMLGKHQIWQQRAREEIINA  360
NES allele2          GHANKKQDINKLNIEEIVNECKLFYLAGQDTTSSLLLWTLIMLGKHQIWQQRAREEIINT  360
NER allele1          GHANKKQDINKLSIEEVVNECKLFYLAGQDTTSSLLLWTLIMLGKHQIWQQRAREEIINA  360
NER allele2          GHANKKQDINKLNIEEIVNECKLFYLAGQDTTSSLLLWTLIMLGKHQIWQQRAREEIINT  360

CYP72A219_4284 genom FGDAKPQFHELNSLKIVNMILQEVLRLYPPVLELNREVTQDIKIGDVILPAGVLLNIPIL  420
NES allele1          FGDAKPQFHELNSLKIVNMILQEVLRLYPPILELNREVTQDIKIGDVILPAGVLLNIPIL  420
NES allele2          FGDAKPQFHELNSLKIVNMILQEVLRLYPPVLELNREVTQDIKIGDVILPAGVLLNIPIL  420
NER allele1          FGDAKPQFHELNSLKIVNMILQEVLRLYPPILELNREVTQDIKIGDVILPAGVLLNIPIL  420
NER allele2          FGDAKPQFHELNSLKIVNMILQEVLRLYPPVLELNREVTQDIKIGDVILPAGVLLNIPIL  420
```

```
CYP72A219_4284 genom LLQQDEKIWGKDAKEFNPERFSEGISKASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLV  480
NES allele1          LLQQDEKIWGKDAKEFNPERFSEGISKASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLV  480
NES allele2          LLQQDEKIWGKDAKEFNPERFSEGISKASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLV  480
NER allele1          LLQQDEKIWGKDAKEFNPERFSEGISKASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLV  480
NER allele2          LLQQDEKIWGKDAKEFNPERFSEGISKASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLV  480

CYP72A219_4284 genom FILQRFVFELSPSYRHAPLNVGTLRPQFGAPIIFRRR                          517
NES allele1          FILQRFVFELSPSYRHAPLNVGTLRPQFGAPIIFRRR                          517
NES allele2          FILQRFVFELSPSYRHAPLNVGTLRPQFGAPIIFRRR                          517
NER allele1          FILQRFVFELSPSYRHAPLNVGTLRPQFGAPIIFRRR                          517
NER allele2          FILQRFVFELSPSYRHAPLNVGTLRPQFGAPIIFRRR                          517
```

CYP72A219_4285

NES and NER CYP72A219_4285 alleles from sanger sequencing
Reference genome use for the alignment:
palmeri(v1.1, unmasked), Name: mRNA4452, Type: CDS, Feature
Location: (Chr: Scaffold_10,
join(13525670..13525754,13526753..13526795,13528419..13528549,13530
643..13530713,13530833..13530895,13531617..13531861,13533902..13534
295,13535370..13535807)) Genomic Location: 13525670-13535807.

>CYP72A219_4285 NES allele1
ATGGAAGAGAAAAGCATATTATCAATTGGAATATCTTTGTTTTGTGTAATATTATTTAGTGTGTCATGGAAAATTATAAATTGGGT
ATGGTTGAAACCAAAAAAGTTCGAAAAATTGCTTAAAAATCAAGGGTTTGATGGCAATTCTTATAGGTTATTGCATGGAGACACC
AAAGATAAAGCTAAGATGACCATGGAAGTTACTTCTAAGCCTATGCCTAATCTATCCAATGATTTTCTTCCACATACTCTTGCAT T
TTTTCATCACACTATATCTAAGTTCGGAAGAAAATGCATGATATGGAATGGGCCGGTTCCATTTGCAAGCATAGCTGAACCGGATT
TGGTTCGGGAAGTTTTTATGAAGATTAACCAATTTCAAAATCCTAAGTTAAACCCTATCGCAAACAAGGTTATATCCGGTTTGATT
ACCTTGGAGGGGAATGCATGGACCCACCGAAGGAAACTCCTTACCCCTGCTTTCCACTTCGATAAACTCAAGTTTATGGTTCC AGC
ATTATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGTC
ATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCAGCTTTTGGAAGTAGTTATCAAGAAGGGAAGACAATTTTTGATCTTA
TAACTGAACAAATTAAAATTGCAATTCCAATGTTTAATTCTGTTTATATCCCCGGTTCGAGGTTTTTGCCAACCAAGAGAAAT AGA
AGAATAACACAAATTGATAGAGAAATCAAGAATCTACTAATGGGTATAATTGACAAAAGAAAGAAGGCAATAAAAACTCAAGAA
AAAACAAAGGATGATTTGCTTAATACATTATTGGAGTCAAGTTATGAGCAAGAAGAAAATGGGCATGGAAATAAAAAAAATGAT
ATTAATTTAAATATTGAAGAAATTCTTGGTGAATGTAAATTATTTTACCTAGCAGGACAAGATACAACTGCTTCTTTATTGGTAT G
GACATTAATTATGTTGGCAAAGCATCAAAATTGGCAACAACTAGCAAGACAAGAGATTTTCAATGCATTTGGGGATAAAACACCT
CAATTTCATGAATTGAACCATCTTAAAATAGTAAATATGATACTACAAGAGGTGCTAAGATTATATCCACCAGTAACAGAAACAA
ATCGTGAAGTAATCCAAGACATAAAATTAGGAAATATGATACTTCCAGCAGGTGTCGTAGTGAATATTCCAATTCTACTTATGCAC
CATGATGAAAAGATATGGGGTAATGATGCGAAAGAATTTAATCCAGAAAGATTTAGCGAAGGAATATCAAAGGCTAGTAAAGGA
AATATGTCATTATTTGCGTTTGGTTGGGGACCTAGGATATGTATTGGATCTAACTTTGCTATGATTGAGGCCAAACTGACATTGGT
TTGCATTTTACAAAGGTTTGTATTTGAGCTTTCACCATCTTATCTTCACGCCCCTTCTGCTATTGGAAGTCTTAGACCCCAATTTGGT
GCTCCAATCATATTTCGTCCACGAATCTTGGGATCATAA

>CYP72A219_4285 NES allele2
ATGGAAGAGAAAAGCATATTATCAATTGGAATATCTTTGTTTTGTGTAATATTATTTAGTGTGTCATGGAAAATTATAAATTGGGT
ATGGTTAAAACCAAAAAAGTTTGAAAAATTGCTTAAAAATCAAGGGTTTGATGGCAATTCTTATAGGTTATTGCATGGAGACACC
AAAGATAAAGCTAAGATGACCATGGAAGTTACTTCTAAGCCTATGCCTAATCTATCCAATGATTTTCTTCCACATACTCTTGCAT T
TTTTCATCACACTATATCTAAGTTCGGAAGAAAATGCATGATATGGAATGGACCGGTTCCATTTGCAAGCATAGCTGAACCGGATT
TGGTCCGGGAAGTTTTTATGAAGATTAACCAATTTCAAAAGCCTAAGTTAAACCCTACCGCAAACAAGGTTATATCCGGTTTGATT
ACCTTGGAGGGGAATGCATGGACCCACCGAAGGAAACTCCTTACCCCTGCTTTCCATTTCGATAAACTCAAGTTTATGGTTCC AGC
ATTATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATTGAAGTTTGGTC
ATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCAGCTTTTGGAAGTAGTTATGAAGAAGGGAAGACAATTTTTGATCTTA
TAACTGAACAAATTAAAATTGCAATTCCAATGTTTAATTCTGTTTATATCCCCGGTTCGAGGTTTTTGCCAACCAAGAGAAAT AGA
AGAATAACACAAATTGATAGAGAAATCAAGAATCTACTAATGGGTATTATTGACAAAAGAAAGAAGGCAATAAAAACTCAAGAA
AAGACAAAGGATGATTTGCTTAATACATTATTGGAGTCAAGTTATGAGCAAGAAGAAAATGGGCATGGAAATAAAAAGAATAAT
ATTAATTTAAATATTGAAGAAATTCTTGGTGAATGTAAATTATTTTACCTAGCAGGACAAGATACAACTGCTTCTTTATTGGTAT G
GACATTAATTATGTTGGCAAAGCATCAAAATTGGCAACAACTAGCAAGACAAGAGATTTTCAATGCATTTGGGGATAAAATACCT
CAATTTCATGAATTGAACCATCTTAAAATAGTAAATATGATACTACAAGAGGTGCTAAGATTATATCCACCAGTAACAGAAACAA
ATCGTGAAGTAATCCAAGACATAAAATTAGGAAATATGATACTTCCAGCAGGTGTCGTAGTGAATATTCCAATTCTACTTATGCAC
CATGATGAAAAGATATGGGGTAATGATGCGAAAGAATTTAATCCAGAAAGATTTAGCGAAGGAATATCAAAGGCTAGTAAAGGA
AATATGTCATTATTTGCGTTTGGTTGGGGACCTAGGATATGTATTGGATCTAACTTTGCTATGATTGAGGCCAAACTGACATTGGT
TTGCATTTTACAAAGGTTTGTATTTGAGCTTTCACCATCTTATCTTCACGCCCCTTCTGCTATTGGAAGTCTTAGACCCCAATTTGGT
GCTCCAATCATATTTCGTCCACGAATCTTGGGATCATAA

>CYP72A219_4285 NER allele1
ATGGAAGAGAAAAGCATATTATCAATTGGAATATCTTTGTTTTGTGTAATATTATTTAGTGTGTCATGGAAAATTATAAATTGGGT
ATGGTTGAAACCAAAAAAGTTCGAAAAATTGCTTAAAAATCAAGGGTTTGATGGCAATTCTTATAGGTTATTGCATGGAGACACC
AAAGATAAAGCTAAGATGACCATGGAAGTTACTTCTAAGCCTATGCCTAATCTATCCAATGATTTTCTTCCACATACTCTTGCAT T
TTTTCATCACACTATATCTAAGTTCGGAAGAAAATGCATGATATGGAATGGGCCGGTTCCATTTGCAAGCATAGCTGAACCGGATT

TGGTTCGGGAAGTTTTTATGAAGATTAACCAATTTCAAAATCCTAAGTTAAACCCTATCGCAAACAAGGTTATATCCGGTTTGATT
ACCTTGGAGGGGAATGCATGGACCCACCGAAGGAAACTCCTTACCCCTGCTTTCCACTTCGATAAACTCAAGTTTATGGTTCCAGC
ATTATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATAGAAGTTTGGTC
ATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCAGCTTTTGGAAGTAGTTATCAAGAAGGGAAGACAATTTTTGATCTTA
TAACTGAACAAATTAAAATTGCAATTCCAATGTTTAATTCTGTTTATATCCCCGGTTCGAGGTTTTTGCCAACCAAGAGAAATAGA
AGAATAACACAAATTGATAGAGAAATCAAGAATCTACTAATGGGTATAATTGACAAAAGAAAGAAGGCAATAAAAACTCAAGAA
AAAACAAAGGATGATTTGCTTAATACATTATTGGAGTCAAGTTATGAGCAAGAAGAAAATGGGCATGGAAATAAAAAAAATGAT
ATTAATTTAAATATTGAAGAAATTCTTGGTGAATGTAAATTATTTTACCTAGCAGGACAAGATACAACTGCTTCTTTATTGGTATG
GACATTAATTATGTTGGCAAAGCATCAAAATTGGCAACAACTAGCAAGACAAGAGATTTTCAATGCATTTGGGGATAAAACACCT
CAATTTCATGAATTGAACCATCTTAAAATAGTAAATATGATACTACAAGAGGTGCTAAGATTATATCCACCAGTAACAGAAACAA
ATCGTGAAGTAATCCAAGACATAAAATTAGGAAATATGATACTTCCAGCAGGTGTCGTAGTGAATATTCCAATTCTACTTATGCAC
CATGATGAAAAGATATGGGGTAATGATGCGAAAGAATTTAATCCAGAAAGATTTAGCGAAGGAATATCAAAGGCTAGTAAAGGA
AATATGTCATTATTTGCGTTTGGTTGGGGACCTAGGATATGTATTGGATCTAACTTTGCTATGATTGAGGCCAAACTGACATTGGT
TTGCATTTTACAAAGGTTTGTATTTGAGCTTTCACCATCTTATCTTCACGCCCCTTCTGCTATTGGAAGTCTTAGACCCCAATTTGGT
GCTCCAATCATATTTCGTCCACGAATCTTGGGATCATAA

>CYP72A219_4285 NER allele2
ATGGAAGAGAAAAGCATATTATCAATTGGAATATCTTTGTTTTGTGTAATATTATTTAGTGTGTCATGGAAAATTATAAATTGGGT
ATGGTTAAAACCAAAAAAGTTTGAAAAATTGCTTAAAAATCAAGGGTTTGATGGCAATTCTTATAGGTTATTGCATGGAGACACC
AAAGATAAAGCTAAGATGACCATGGAAGTTACTTCTAAGCCTATGCCTAATCTATCCAATGATTTTCTTCCACATACTCTTGCATT
TTTTCATCACACTATATCTAAGTTCGGAAGAAAATGCATGATATGGAATGGACCGGTTCCATTTGCAAGCATAGCTGAACCGGATT
TGGTCCGGGAAGTTTTTATGAAGATTAACCAATTTCAAAAGCCTAAGTTAAACCCTACCGCAAACAAGGTTATATCCGGTTTGATT
ACCTTGGAGGGGAATGCATGGACCCACCGAAGGAAACTCCTTACCCCTGCTTTCCATTTCGATAAACTCAAGTTTATGGTTCCAGC
ATTATGGGAAAGTTCAAAAGAAATGATTAAAGAATGGGAAGAATTGGCATCAAAAACAGGTTCATGTGAAATTGAAGTTTGGTC
ATCTTTACAAAAATTATCAGCAGATATGATTTCTAGAGCAGCTTTTGGAAGTAGTTATGAAGAAGGGAAGACAATTTTTGATCTTA
TAACTGAACAAATTAAAATTGCAATTCCAATGTTTAATTCTGTTTATATCCCCGGTTCGAGGTTTTTGCCAACCAAGAGAAATAGA
AGAATAACACAAATTGATAGAGAAATCAAGAATCTACTAATGGGTATTATTGACAAAAGAAAGAAGGCAATAAAAACTCAAGAA
AAGACAAAGGATGATTTGCTTAATACATTATTGGAGTCAAGTTATGAGCAAGAAGAAAATGGGCATGGAAATAAAAAGAATAAT
ATTAATTTAAATATTGAAGAAATTCTTGGTGAATGTAAATTATTTTACCTAGCAGGACAAGATACAACTGCTTCTTTATTGGTATG
GACATTAATTATGTTGGCAAAGCATCAAAATTGGCAACAACTAGCAAGACAAGAGATTTTCAATGCATTTGGGGATAAAATACCT
CAATTTCATGAATTGAACCATCTTAAAATAGTAAATATGATACTACAAGAGGTGCTAAGATTATATCCACCAGTAACAGAAACAA
ATCGTGAAGTAATCCAAGACATAAAATTAGGAAATATGATACTTCCAGCAGGTGTCGTAGTGAATATTCCAATTCTACTTATGCAC
CATGATGAAAAGATATGGGGTAATGATGCGAAAGAATTTAATCCAGAAAGATTTAGCGAAGGAATATCAAAGGCTAGTAAAGGA
AATATGTCATTATTTGCGTTTGGTTGGGGACCTAGGATATGTATTGGATCTAACTTTGCTATGATTGAGGCCAAACTGACATTGGT
TTGCATTTTACAAAGGTTTGTATTTGAGCTTTCACCATCTTATCTTCACGCCCCTTCTGCTATTGGAAGTCTTAGACCCCAATTTGGT
GCTCCAATCATATTTCGTCCACGAATCTTGGGATCATAA

CLUSTAL alignment

```
CYP72A219_4285 genom ------------MCFLAQWFSWEGGDLPSMTDDVSDDISFHSEIHAFSAKILYNTLLHG   47
NES allele1          MEEKSILSIGISLFCVILFSVSWKIINWVWLKPK-----KFEKLLKNQGFDGNSYRLLHG   55
NES allele 2         MEEKSILSIGISLFCVILFSVSWKIINWVWLKPK-----KFEKLLKNQGFDGNSYRLLHG   55
NER allele 1         MEEKSILSIGISLFCVILFSVSWKIINWVWLKPK-----KFEKLLKNQGFDGNSYRLLHG   55
NER allele 2         MEEKSILSIGISLFCVILFSVSWKIINWVWLKPK-----KFEKLLKNQGFDGNSYRLLHG   55

CYP72A219_4285 genom DTKDKAKMTMEVTSKPMPNLSNDFLPHTLAFFHHTISKFGRKCIIWNGPVPFASIAEPDL  107
NES allele1          DTKDKAKMTMEVTSKPMPNLSNDFLPHTLAFFHHTISKFGRKCMIWNGPVPFASIAEPDL  115
NES allele 2         DTKDKAKMTMEVTSKPMPNLSNDFLPHTLAFFHHTISKFGRKCMIWNGPVPFASIAEPDL  115
NER allele 1         DTKDKAKMTMEVTSKPMPNLSNDFLPHTLAFFHHTISKFGRKCMIWNGPVPFASIAEPDL  115
NER allele 2         DTKDKAKMTMEVTSKPMPNLSNDFLPHTLAFFHHTISKFGRKCMIWNGPVPFASIAEPDL  115

CYP72A219_4285 genom VR--------------------------EGNAWTHRRKLLTPAFHFDKLKFMVPALWE  139
NES allele1          VREVFMKINQFQNPKLNPIANKVISGLITLEGNAWTHRRKLLTPAFHFDKLKFMVPALWE  175
NES allele 2         VREVFMKINQFQKPKLNPTANKVISGLITLEGNAWTHRRKLLTPAFHFDKLKFMVPALWE  175
NER allele 1         VREVFMKINQFQNPKLNPIANKVISGLITLEGNAWTHRRKLLTPAFHFDKLKFMVPALWE  175
NER allele 2         VREVFMKINQFQKPKLNPTANKVISGLITLEGNAWTHRRKLLTPAFHFDKLKFMVPALWE  175

CYP72A219_4285 genom STKEMIEEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYQEGKTIFDLITEQIKI  199
NES allele1          SSKEMIKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYQEGKTIFDLITEQIKI  235
NES allele 2         SSKEMIKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYEEGKTIFDLITEQIKI  235
NER allele 1         SSKEMIKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYQEGKTIFDLITEQIKI  235
NER allele 2         SSKEMIKEWEELASKTGSCEIEVWSSLQKLSADMISRAAFGSSYEEGKTIFDLITEQIKI  235

CYP72A219_4285 genom AIPMFNSVYIPGSRFLPTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEKTKDDLLNTL  259
NES allele1          AIPMFNSVYIPGSRFLPTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEKTKDDLLNTL  295
NES allele 2         AIPMFNSVYIPGSRFLPTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEKTKDDLLNTL  295
NER allele 1         AIPMFNSVYIPGSRFLPTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEKTKDDLLNTL  295
NER allele 2         AIPMFNSVYIPGSRFLPTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEKTKDDLLNTL  295
```

```
CYP72A219_4285 genom LESSYEQEENGHGNKKNDINLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQNWQQ  319
NES allele1          LESSYEQEENGHGNKKNDINLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQNWQQ  355
NES allele 2         LESSYEQEENGHGNKKNNINLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQNWQQ  355
NER allele 1         LESSYEQEENGHGNKKNDINLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQNWQQ  355
NER allele 2         LESSYEQEENGHGNKKNNINLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQNWQQ  355

CYP72A219_4285 genom LARQEIFNAFGDKIPQFHDLNHLKIVNMILQEVLRLYPPVTETNREVIQDIKLGNMILPA  379
NES allele1          LARQEIFNAFGDKTPQFHELNHLKIVNMILQEVLRLYPPVTETNREVIQDIKLGNMILPA  415
NES allele 2         LARQEIFNAFGDKIPQFHELNHLKIVNMILQEVLRLYPPVTETNREVIQDIKLGNMILPA  415
NER allele 1         LARQEIFNAFGDKTPQFHELNHLKIVNMILQEVLRLYPPVTETNREVIQDIKLGNMILPA  415
NER allele 2         LARQEIFNAFGDKIPQFHELNHLKIVNMILQEVLRLYPPVTETNREVIQDIKLGNMILPA  415

CYP72A219_4285 genom GVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGISKASKGNMSLFAFGWGPRICIGSNFA  439
NES allele1          GVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGISKASKGNMSLFAFGWGPRICIGSNFA  475
NES allele 2         GVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGISKASKGNMSLFAFGWGPRICIGSNFA  475
NER allele 1         GVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGISKASKGNMSLFAFGWGPRICIGSNFA  475
NER allele 2         GVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGISKASKGNMSLFAFGWGPRICIGSNFA  475

CYP72A219_4285 genom MIEAKLTLVCILQRFVFELSPSYLHAPSAIGSLRPQFGAPIIFRPRILGS            489
NES allele1          MIEAKLTLVCILQRFVFELSPSYLHAPSAIGSLRPQFGAPIIFRPRILGS            525
NES allele 2         MIEAKLTLVCILQRFVFELSPSYLHAPSAIGSLRPQFGAPIIFRPRILGS            525
NER allele 1         MIEAKLTLVCILQRFVFELSPSYLHAPSAIGSLRPQFGAPIIFRPRILGS            525
NER allele 2         MIEAKLTLVCILQRFVFELSPSYLHAPSAIGSLRPQFGAPIIFRPRILGS            525
```

CYP72A219_7285

NES and NER CYP72A219_7285 alleles from sanger sequencing
Reference genome use for the alignment:
>Amaranthus palmeri(v1.1, NCBI WindowMasker (Hard)), Name: mRNA7568, Type: CDS, Feature Location: (Chr: Scaffold_12, join(10909788..10909884,10914812..10915029,10916014..10916258,10917355..10917751,10918716..10919153)) Genomic Location: 10909788-10919153

>CYP72A219_7285 NER allele 1
ATGGTGAAAGAAACTACATCTAAACCCATGGATCATTTTTCAAATGATTATATCCCTAGGGTCTTCCCATTTCATCATCAAATTAT
CAAAAATTATGGTGAAAAATCATTTGTATGGGATGGTCCAATCCCATCAATAATGATAACAAACCCAGAAATAATAAGAGAAATA
TTAATGAAAATTTATGATTTTAAGAAAATAGTGCCAACTCCACTTATGAAGAGGCTCACTAATGGGCTAGTAAGGTTGGAAGGGA
CCAAATGGGCCAAAGAAAGAAAGCTTCTTAATCCTGCATTCCATATGGAGAAATTAAAGCATATGATACCAGCATTTCATGCAAG
TTGTAAAGAAATGATAAAAGAATGGGAGAAAATAATATCAAAAACAGGAAGTGGGGAGGTAGAAATGTGGTCATATCTACATAA
TATGTCAGCAGATGCAATCTCTAGGGCAGCATTTGGAAGTAGTTTTCAAGAAGGCCAAAGGGTTTTTGAACTTCTAAGAGAACAT
ACTTCTATTACTGTACAAGCTTTGCAATCTGTCTATATTCCTGGATCCAGGTTCCTTCCAACCAAAACCAATAGAAGGGAGGCAAA
AATAATCAAAGAAATGGAATCATTACTCAAAGCTATGATCCAAAGAAGGCAAAATGCTATAAAAAATGGAGAAAATGGGAATAA
GGATGACTTATTGGATTTATTATTGACATCAAGCATAGAAGATAATCATGTAAAAGATGGAAATACCAAAGAATATGTTAAATTG
GGTATGAAAGAAATAATTGAAGAATGCAAAATGTTTTACATAGCTGGACAAGAAACCACTTCTTCTTTGTTGACATGGACACTTA
TTTTGTTAAGTAAGCACCAAAATTGGCAACACCAAGCAAGAGAAGAGATTCTTGCCACATTTGGGGATAGTATCCCTAATGATGA
GGGCTTAAACCAATTGAAGAAGGTAAGCATGATACTTCAAGAGGTTCTAAGGTTATATCCACCGGTAGTTACCTTAGGTCGAAAA
GTTTCTCATGACATAAAGGTTGGAGATATGGATTTACCCGCCGGTGTGCAAGTTAAAATTCCAACTATTTTTATTCACCACAGTGA
AAAATTATGGGGTAAAGATGCAAAAGAGTTTAATCCGAATAGATTTTCACAAGGACTTTTGAAAGCCACAGGAGGAAATATGTGT
TTCTTTGCATTCGGATGGGGCCCACGTGTATGTATTGGCTCTAATTTTGCATTACTTGAGGCTAAAATGGCATTAGCTTTGATTTTG
CCGCACTTTTCATTTGAGTTGTCACCGTCCTACGCTCATGCTCCTACCATCGGTCAGGGTACTCTGCGCCCTCAATTTGGTGCTAAG
TTAATTTTACACCGTATAAACCGTATTTGA

>CYP72A219_7285 NER allele 2
ATGGTGAAAGAAACTACATCTAAACCCATGGATCATTTTTCAAATGATTATATCCCTAGGGTCTTCCCATTTCATCATCAAATTAT
CAAAAATTATGGTGAAAAATCATTTGTATGGGATGGTCCAATTCCATCAATAATGATAACAAACCCAGAAGAAATAAGAATAATA
TTAATGAAAATTTATGATTTTAAGAAAATAGTGCCAACTCCACTTATGAAGAGGCTCACTAATGGGCTAGTAAGGTTGGAAGGGA
CCAAATGGGCCAAAGAAAGAAAGCTTCTTAATCCTGCCTTCCATATGGAGAAATTAAAGCATATGATACCAGCATTTCATGCAAG
TTGTAAAGAAATGATAAAAGAATGGGAGAAAATAATATCAAAAACAGGAAATGGAGAGGTAGAAATGTGGTCATATCTACATAA
TATGTCAGCAGATGCAATCTCTAGGGCAGCATTTGGAAGTAGTTTTCAAGAAGGCCAAAGGGTTTTTGAACTTCTAAGAGAACAT
ACTTCTATTACTGTACAAGCTTTGCAATCTGTCTATATTCCTGGATCCAGGTTCCTTCCAACCAAAACCAATAGAAGGGAGGCAAA
AATAATCAAAGAAATGGAATCATTACTCAAAGCTATGATCCAAAGAAGGCAAAATGCTATAAAAAAGGGAGAAAATGGGAATAA
GGATGACTTATTGGATTTATTATTGACATCAAGCATAGAAGATAATCATGTAAAAGATGGAAATACCAAAGAATATGTTAAATTG
GGTATGAAAGAAATAATTGAAGAATGCAAAATGTTTTACATAGCTGGACAAGAAACCACTTCTTCTTTGTTGACATGGACACTTA
TTTTGTTAAGTAAGCACCAAAATTGGCAACACCAAGCAAGAGAAGAGATTCTAGCCACATTTGGGGATACTATCCCTAATGATGA
GGGCTTAAACCAATTGAAGAAGGTAAGCATGATACTTCAAGAGGTTCTGAGGTTATATCCACCGGTAGTTACCTTAGGTCGAAAA
GTTTCTCATGACATAAAGGTTGGAGATATGGATTTACCAGCCGGTGTGCAAGTTAAAATTCCAACTATTTTTATTCACCACAGTGA
AAAACTATGGGGTAAAGATGCAAAAGAGTTTAACCCGAATAGATTTTCACAAGGATTTTTGAAAGCCACAGGAGGAAATATGTGT
TTCTTTGCATTCGGATGGGGCCCACGTATATGTATTGGCTCTAATTTTGCATTACTTGAGGCTAAAATGGCATTAGCTTTGATTTTG
CCGCACTTTTCATTTGAGTTGTCACCGTCCTACGCTCATGCTCCTACAATCGGTCAGGGTACTCTGCGCCCTCAATTTGGTGCTAAG
TTAATTTTACACCGTATAAACCGTATTTGA

> CYP72A219_7285 NES allele 1

ATGGTGAAAGAAACTACATCTAAACCCATGGATCATTTTTCAAATGATTATATCCCTAGGGTCTTCCCATTTCATCATCAAATTAT
CAAAAATTATGGTGAAAAATCATTTGTATGGGATGGTCCAATTCCATCAATAATGATAACAAACCCAGAAGAAATAAGAATAATA
TTAATGAAAATTTATGATTTTAAGAAAATAGTGCCAACTCCACTTATGAAGAGGCTCACTAATGGGCTAGTAAGGTTGGAAGGGA
CCAAATGGGCCAAAGAAAGAAAGCTTCTTAATCCTGCCTTCCATATGGAGAAATTAAAGCATATGATACCAGCATTTCATGCAAG
TTGTAAAGAAATGATAAAAGAATGGGAGAAAATAATATCAAAAACAGGAAATGGAGAGGTAGAAATGTGGTCATATCTACATAA
TATGTCAGCAGATGCAATCTCTAGGGCAGCATTTGGAAGTAGTTTTCAAGAAGGCCAAAGGGTTTTTGAACTTCTAAGAGAACAT
ACTTCTATTACTGTACAAGCTTTGCAATCTGTCTATATTCCTGGATCCAGGTTCCTTCCAACCAAAACCAATAGAAGGGAGGCAAA
AATAATCAAAGAAATGGAATCATTACTCAAAGCTATGATCCAAAGAAGGCAAAATGCTATAAAAAAGGGAGAAAATGGGAATAA
GGATGACTTATTGGATTTATTATTGACATCAAGCATAGAAGATAATCATGTAAAAGATGGAAATACCAAAGAATATGTTAAATTG
GGTATGAAAGAAATAATTGAAGAATGCAAAATGTTTTACATAGCTGGACAAGAAACCACTTCTTCTTTGTTGACATGGACACTTA
TTTTGTTAAGTAAGCACCAAAATTGGCAACACCAAGCAAGAGAAGAGATTCTAGCCACATTTGGGGATACTATCCCTAATGATGA
GGGCTTAAACCAATTGAAGAAGGTAAGCATGATACTTCAAGAGGTTCTGAGGTTATATCCACCGGTAGTTACCTTAGGTCGAAAA
GTTTCTCATGACATAAAGGTTGGAGATATGGATTTACCAGCCGGTGTGCAAGTTAAAATTCCAACTATTTTTATTCACCACAGTGA
AAAACTATGGGGTAAAGATGCAAAAGAGTTTAACCCGAATAGATTTTCACAAGGATTTTTGAAAGCCACAGGAGGAAATATGTGT
TTCTTTGCATTCGGATGGGGCCCACGTATATGTATTGGCTCTAATTTTGCATTACTTGAGGCTAAAATGGCATTAGCTTTGATTTTG
CCGCACTTTTCATTTGAGTTGTCACCGTCCTACGCTCATGCTCCTACAATCGGTCAGGGTACTCTGCGCCCTCAATTTGGTGCTAAG
TTAATTTTACACCGTATAAACCGTATTTGA

> CYP72A219_7285 NES allele 2

ATGGTGAAAGAAACTACATCTAAACCCATGGATCATTTTTCAAATGATTATATCCCTAGGGTCTTCCCATTTCATCATCAAATTAT
CAAAAATTATGGTGAAAAATCATTTGTATGGGATGGTCCAATTCCATCAATAATGATAACAAACCCAGAAGAAATAAGAATAATA
TTAATGAAAATTTATGATTTTAAGAAAATAGTGCCAACTCCACTTATGAAGAGGCTCACTAATGGGCTAGTAAGGTTGGAAGGGA
CCAAATGGGCCAAAGAAAGAAAGCTTCTTAATCCTGCCTTCCATATGGAGAAATTAAAGCATATGATACCAGCATTTCATGCAAG
TTGTAAAGAAATGATAAAAGAATGGGAGAAAATAATATCAAAAACAGGAAATGGAGAGGTAGAAATGTGGTCATATCTACATAA
TATGTCAGCAGATGCAATCTCTAGGGCAGCATTTGGAAGTAGTTTTCAAGAAGGCCAAAGGGTTTTTGAACTTCTAAGAGAACAT
ACTTCTATTACTGTACAAGCTTTGCAATCTGTCTATATTCCTGGATCCAGGTTCCTTCCAACCAAAACCAATAGAAGGGAGGCAAA
AATAATCAAAGAAATGGAATCATTACTCAAAGCTATGATCCAAAGAAGGCAAAATGCTATAAAAAAGGGAGAAAATGGGAATAA
GGATGACTTATTGGATTTATTATTGACATCAAGCATAGAAGATAATCATGTAAAAGATGGAAATACCAAAGAATATGTTAAATTG
GGTATGAAAGAAATAATTGAAGAATGCAAAATGTTTTACATAGCTGGACAAGAAACCACTTCTTCTTTGTTGACATGGACACTTA
TTTTGTTAAGTAAGCACCAAAATTGGCAACACCAAGCAAGAGAAGAGATTCTAGCCACATTTGGGGATACTATCCCTAATGATGA
GGGCTTAAACCAATTGAAGAAGGTAAGCATGATACTTCAAGAGGTTCTGAGGTTATATCCACCGGTAGTTACCTTAGGTCGAAAA
GTTTCTCATGACATAAAGGTTGGAGATATGGATTTACCAGCCGGTGTGCAAGTTAAAATTCCAACTATTTTTATTCACCACAGTGA
AAAACTATGGGGTAAAGATGCAAAAGAGTTTAACCCGAATAGATTTTCACAAGGATTTTTGAAAGCCACAGGAGGAAATATGTGT
TTCTTTGCATTCGGATGGGGCCCACGTATATGTATTGGCTCTAATTTTGCATTACTTGAGGCTAAAATGGCATTAGCTTTGATTTTG
CCGCACTTTTCATTTGAGTTGTCACCGTCCTACGCTCATGCTCCTACAATCGGTCAGGGTACTCTGCGCCCTCAATTTGGTGCTAAG
TTAATTTTACACCGTATAAACCGTATTTGA

CLUSTAL alignment

```
CYP72A219_7285 genome   MVKETTSKPMDHFSNDYIPRVFPFHHQIIKNYGEKSFVWDGPIPSIMITNPEIIREILMK   60
NES allele1             MVKETTSKPMDHFSNDYIPRVFPFHHQIIKNYGEKSFVWDGPIPSIMITNPEEIRIILMK   60
NES allele2             MVKETTSKPMDHFSNDYIPRVFPFHHQIIKNYGEKSFVWDGPIPSIMITNPEEIRIILMK   60
NER allele1             MVKETTSKPMDHFSNDYIPRVFPFHHQIIKNYGEKSFVWDGPIPSIMITNPEEIRIILMK   60
NER allele2             MVKETTSKPMDHFSNDYIPRVFPFHHQIIKNYGEKSFVWDGPIPSIMITNPEIIREILMK   60

CYP72A219_7285 genome   IYDFKKIVPTPLMKRLTNGLVRLEGTKWAKERKLLNPAFHMEKLKHMIPAFHASCKEMIK  120
NES allele1             IYDFKKIVPTPLMKRLTNGLVRLEGTKWAKERKLLNPAFHMEKLKHMIPAFHASCKEMIK  120
NES allele2             IYDFKKIVPTPLMKRLTNGLVRLEGTKWAKERKLLNPAFHMEKLKHMIPAFHASCKEMIK  120
NER allele1             IYDFKKIVPTPLMKRLTNGLVRLEGTKWAKERKLLNPAFHMEKLKHMIPAFHASCKEMIK  120
NER allele2             IYDFKKIVPTPLMKRLTNGLVRLEGTKWAKERKLLNPAFHMEKLKHMIPAFHASCKEMIK  120

CYP72A219_7285 genome   EWEKIISKTGSGEVEMWSYLHNMSADAISRAAFGSSFQEGQRVFELLREHTSITVQALQS  180
NES allele1             EWEKIISKTGNGEVEMWSYLHNMSADAISRAAFGSSFQEGQRVFELLREHTSITVQALQS  180
NES allele2             EWEKIISKTGNGEVEMWSYLHNMSADAISRAAFGSSFQEGQRVFELLREHTSITVQALQS  180
NER allele1             EWEKIISKTGNGEVEMWSYLHNMSADAISRAAFGSSFQEGQRVFELLREHTSITVQALQS  180
NER allele2             EWEKIISKTGSGEVEMWSYLHNMSADAISRAAFGSSFQEGQRVFELLREHTSITVQALQS  180

CYP72A219_7285 genome   VYIPGSRFLPTKTNRREAKIIKEMESLLKAMIQRRQNAIKNGENEHKDDLLDLLLTSSIE  240
NES allele1             VYIPGSRFLPTKTNRREAKIIKEMESLLKAMIQRRQNAIKKGENGNKDDLLDLLLTSSIE  240
NES allele2             VYIPGSRFLPTKTNRREAKIIKEMESLLKAMIQRRQNAIKKGENGNKDDLLDLLLTSSIE  240
NER allele1             VYIPGSRFLPTKTNRREAKIIKEMESLLKAMIQRRQNAIKKGENGNKDDLLDLLLTSSIE  240
NER allele2             VYIPGSRFLPTKTNRREAKIIKEMESLLKAMIQRRQNAIKNGENGNKDDLLDLLLTSSIE  240

CYP72A219_7285 genome   DNHVKDGNNKEYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQNWQHQAREEI  300
NES allele1             DNHVKDGNTKEYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQNWQHQAREEI  300
NES allele2             DNHVKDGNTKEYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQNWQHQAREEI  300
NER allele1             DNHVKDGNTKEYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQNWQHQAREEI  300
```

```
NER allele2             DNHVKDGNTKEYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQNWQHQAREEI  300

CYP72A219_7285 genome   IATFGDTIPNDEGLNQLKKVSMILQEVLRLYPPVVTLGRKVSHDIKVGDMDLPAGVQVKI  360
NES allele1             LATFGDTIPNDEGLNQLKKVSMILQEVLRLYPPVVTLGRKVSHDIKVGDMDLPAGVQVKI  360
NES allele2             LATFGDTIPNDEGLNQLKKVSMILQEVLRLYPPVVTLGRKVSHDIKVGDMDLPAGVQVKI  360
NER allele1             LATFGDTIPNDEGLNQLKKVSMILQEVLRLYPPVVTLGRKVSHDIKVGDMDLPAGVQVKI  360
NER allele2             LATFGDSIPNDEGLNQLKKVSMILQEVLRLYPPVVTLGRKVSHDIKVGDMDLPAGVQVKI  360

CYP72A219_7285 genome   PTIFIHHSEKLWGKDAKEFNPNRFSQGFLKATGGNMCFFAFGWGPRICIGSNFALLEAKM  420
NES allele1             PTIFIHHSEKLWGKDAKEFNPNRFSQGFLKATGGNMCFFAFGWGPRICIGSNFALLEAKM  420
NES allele2             PTIFIHHSEKLWGKDAKEFNPNRFSQGFLKATGGNMCFFAFGWGPRICIGSNFALLEAKM  420
NER allele1             PTIFIHHSEKLWGKDAKEFNPNRFSQGFLKATGGNMCFFAFGWGPRICIGSNFALLEAKM  420
NER allele2             PTIFIHHSEKLWGKDAKEFNPNRFSQGLLKATGGNMCFFAFGWGPRVCIGSNFALLEAKM  420

CYP72A219_7285 genome   ALALILPHFSFELSPSYAHAPTIGQGTLRPQFGAKLILHRINRI                  464
NES allele1             ALALILPHFSFELSPSYAHAPTIGQGTLRPQFGAKLILHRINRI                  464
NES allele2             ALALILPHFSFELSPSYAHAPTIGQGTLRPQFGAKLILHRINRI                  464
NER allele1             ALALILPHFSFELSPSYAHAPTIGQGTLRPQFGAKLILHRINRI                  464
NER allele2             ALALILPHFSFELSPSYAHAPTIGQGTLRPQFGAKLILHRINRI                  464
```

CYP81E8

NES and NER CYP81E8 alleles from sanger sequencing
Reference genome use for the alignment:
palmeri(v1.1, unmasked), Name: mRNA14137, Type: CDS, Feature Location: (Chr: Scaffold_2, complement(join(10072920..10073540,10074091..10074957))) Genomic Location: 10072920-10074957.

>NES_CYP81E_allele1
ATGGAAATTTTATTCACTTTTTTATCCATCTTTGCTACTTTTTTCTTCCTCTATAAAATCATCCAATCCCTTAAAACAAACAATAAAA
AACACCCACCAAGCCCACCATCATACCCATTATTGGGCCACTTACACCTTCTAAAACCCCCATTCCACCGTACACTCCAATCTCTA
GCCCAACGTTACGGCCCAATCTTCACTCTCAAATTGGGCCTTCAACGGGCCGTGGTAGTTTCCTCCGCATGGGCCGCCGAAGAATG
TTTCGGTCAAAACGACGTCGTTTTTGCAAACAGACCAAAATTCATAATAGGACAACATTTAGGATACAACCACTCTATCCTTATCT
GGTCCCCATATGGGGACTACTGGCGGAACCTCCGACGTGTCACAACCATTACCATGTTATCCATGAGACGGATCAACGAAGCGGG
TCCGACCCGGAAACTTGAGATCCGAAACATGATAATGAACCTTCTAGAAAGTGGAAGTGGGACCCAAAAGGTGAATATGAATGA
AGTGCTGAGTAAACTAGCAAGAAATTTTGTAATGAGGATAGTAAATGGAAAATCATGGGAAAAAATGATAATAAAACCCCCTGA
AAATTTAATGACTATTTGTGATTTTTTACCATTTTTAAGATGGGTAGATTTTAAAGGGATAGAAAAGGATATGAAGGAAAAACAG
GTTGAAAGAGATGAGTTTTTACAAAGTTTGGTTGATGAATTTAGAGAAAGTAAGAAGAAAGGTGAGGATTTAGGGATAAATAAT
ACTTTGATTGAACAATTGCTGGATTTGCAACAAGCTGAACCTGATCAATACTCTGATGAAACTATCAAAGGAATCATTTTGGTAAT
GTTACTAGCAGGATCAGAAACCACAGCACGAACCTTAGAATGGGCATTATCAAATCTCATAAACCACCCAGAAATCCTCGTGAAA
GCAAGAACAGAAATCGACCTCCATGTAGGCAAGGAACGACTAGTAGATGATTCTGATCTACCAAAATTAACGTATACACGTTGCA
TTGTATACGAAACTCTAAGACTATTTCCGGCTGCACCACTTTTAGTACCTCATTTTTCTTCGCAAGATTGCACCATAGGAGGCTATC
ACGTACCAAAAGGGACCATGCTTTTCGTAAATGCTTGGGCTATACACAGAGACCCGACCCTATGGGACGAGCCTACTGTGTTTAA
GCCTGAGAGGTTCGAGAAGGAGAAAGAAGGGTTTAAGTTTATGCCCTTTGGAATTGGAAGGAGATCTTGTCCAGGGAACAATGTG
GCTCTTAGGAACGTTTCATTGACATTGGCTACTCTTATTCAGTGCTTTGATTGGGAAGCAGCTGAGAGTGGATCAATTGATCTAAC
CGAAAAGCGTGGTGTTGGTGCGATCATTGTCCCTAAGGAAAAGCCATTGGAGGCTATATGTCGACCACGACCCTCAATGGAGGAT
TTCCTCGCTAAAATTTAG

>NES_CYP81E_allele2
ATGGAAATTTTATTCACTTTTTTATCCATCTTTGCTACTTTTTTCTTCCTCTATAAAATCATCCATTCCCTTAAAACAAACAACAAAA
AACACCCACCAAGCCCACCAAGCCCACCATCATACCCATTATTGGGCCACTTACACCTTCTAAAACCCCCATTCCACCGTACACTC
CAATCTCTAGCCCAACGTTACGGCCCAATCTTCACTCTCAAATTGGGCCTTCAACGGGCCGTGGTAGTTTCCTCCGCATGGGCCGC
CGAAGAATGTTTCGGCCAAAACGACGTCGTTTTTGCAAACAGACCAAAATTCATAATAGGACAACATTTAGGATACAACCACTCT
ATCCTTATCTGGTCCCCATACGGCGACTACTGGCGGAACCTCCGACGTGTCACAACCATTACCATGTTATCCATGAGACGGATCAA
CGAAGCGGGTCCGACCCGGAAACTTGAGATCCGAAACATGATAAGGAACCTTCTAGAAAGTGGGAGTGGGACCCAAAAGGTGAA
TATGAATGAAGTGCTGAGTAAACTAGCAAGAAATTTTGTAATGAGGATAGTAAATGGAAAATCATGGGAAAAAATGATAATAAA
ACCCCCTGAAAATTTAATGACTATTTGTGATTTTTTACCATTTTTAAGATGGGTAGATTTTAAAGGGATAGAAAAGGATATGAAGG
AAAAACAGGTTGAAAGAGATGAGTTTTTACAAAGTTTGGTTGATGAATTTAGAGAAAGTAAGAAGAAAGGTGAGGATTTAGGGA
TGAATAATACTTTGATTGAACAATTGCTGGATTTGCAACAAGCTGAACCTGATCAATACTCTGATGAAACTATCAAAGGGATCATT
TTGGTAATGTTACTAGCGGGATCAGAAACCACAGCACGAACCTTAGAATGGGCATTATCAAATCTCATTAACCACCCAGAAATCC
TCGCGAAAGCAAGAACAGAAATCGACCTCCATGTAGGCAAGGAAAGACTAGTAGATGATTCTGATCTCCCAAAATTAACGTATAC
ACGTTGCATTGTATACGAAACTCTAAGACTATTTCCGGCTGCACCACTTTTAGTACCTCATTTTTCTTCACAAGATTGCACCATAGG
AGGCTATCACGTACCAAAAGGGACAATGCTTTTCGTAAATGCTTGGGCTATACACAGAGACCCGACCCTATGGGACGAGCCTACT
TTGTTTAAGCCCGAGAGGTTCGAAAAGGAGAAAGAAGGGTTTAAGTTTATGCCCTTTGGAATTGGAAGGAGATCTTGTCCAGGGA
ACAATGTGGCTCTTAGGAACGTTTCATTGACATTGGCTACTCTTATTCAGTGCTTTGATTGGGAAGCAGCTGAGAGTGGATCAATT
GATCTAACCGAAAAGCGTGGTGTTGGTGCGATCATTGTCCCTAAGGAAAAGCCATTGGAGGCTATATGTCGACCTCGACCCTCAA
TGGAGGATTTCCTCGCTAAAATTTAG

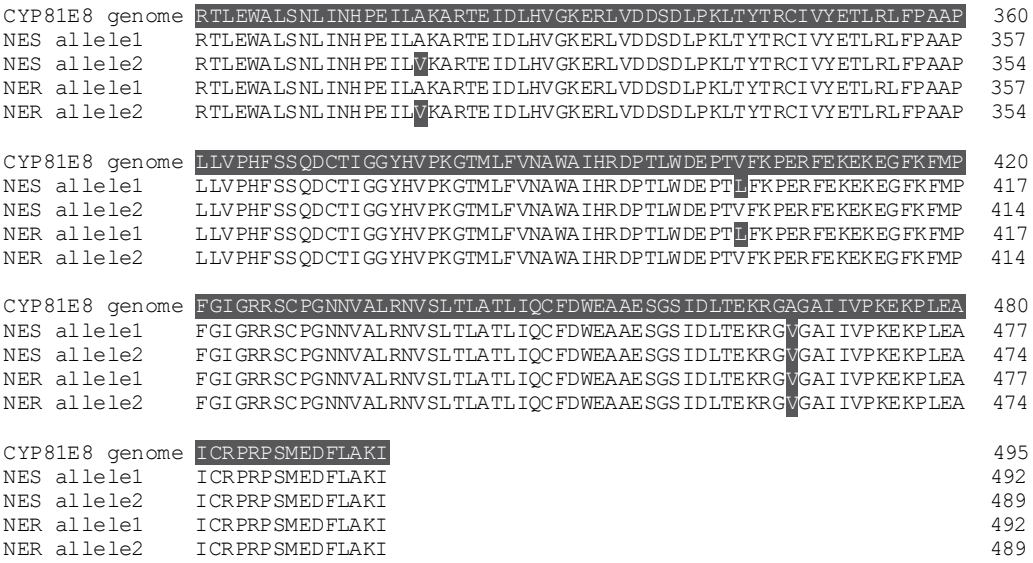
>NER_CYP81E_allele1

ATGGAAATTTTATTCACTTTTTTATCCATCTTTGCTACTTTTTTCTTCCTCTATAAAATCATCCAATCCCTTAAAACAAACAATAAAA
AACACCCACCAAGCCCACCATCATACCCATTATTGGGCCACTTACACCTTCTAAAACCCCCATTCCACCGTACACTCCAATCTCTA
GCCCAACGTTACGGCCCAATCTTCACTCTCAAATTGGGCCTTCAACGGGCCGTGGTAGTTTCCTCCGCATGGGCCGCCGAAG AATG
TTTCGGTCAAAACGACGTCGTTTTTGCAAACAGACCAAAATTCATAATAGGACAACATTTAGGATACAACCACTCTATCCTTATCT
GGTCCCCATATGGGGACTACTGGCGGAACCTCCGACGTGTCACAACCATTACCATGTTATCCATGAGACGGATCAACGAAGCGGG
TCCGACCCGGAAACTTGAGATCCGAAACATGATAATGAACCTTCTAGAAAGTGGAAGTGGGACCCAAAAGGTGAATATGAA TGA
AGTGCTGAGTAAACTAGCAAGAAATTTTGTAATGAGGATAGTAAATGGAAAATCATGGGAAAAAATGATAATAAAACCCCCTGA
AAATTTAATGACTATTTGTGATTTTTTACCATTTTTAAGATGGGTAGATTTTAAAGGGATAGAAAAGGATATGAAGGAAAAACAG
GTTGAAAGAGATGAGTTTTTACAAAGTTTGGTTGATGAATTTAGAGAAAGTAAGAAGAAAGGTGGAGGATTTAGGGATAAATAAT
ACTTTGATTGAACAATTGCTGGATTTGCAACAAGCTGAACCTGATCAATACTCTGATGAAACTATCAAAGGAATCATTTTGGTAAT
GTTACTAGCAGGATCAGAAACCACAGCACGAACCTTAGAATGGGCATTATCAAATCTCATAAACCACCCAGAAATCCTCGTGAAA
GCAAGAACAGAAATCGACCTCCATGTAGGCAAGGAACGACTAGTAGATGATTCTGATCTACCAAAATTAACGTATACACGTTGCA
TTGTATACGAAACTCTAAGACTATTTCCGGCTGCACCACTTTTAGTACCTCATTTTTCTTCGCAAGATTGCACCATAGGAGGCTATC
ACGTACCAAAAGGGACCATGCTTTTCGTAAATGCTTGGGCTATACACAGAGACCCGACCCTATGGGACGAGCCTACTGTGTTTAA
GCCTGAGAGGTTCGAGAAGGAGAAAGAAGGGTTTAAGTTTATGCCCTTTGGAATTGGAAGGAGATCTTGTCCAGGGAACAATGTG
GCTCTTAGGAACGTTTCATTGACATTGGCTACTCTTATTCAGTGCTTTGATTGGGAAGCAGCTGAGAGTGGATCAATTGATCTAAC
CGAAAAGCGTGGTGTTGGTGCGATCATTGTCCCTAAGGAAAAGCCATTGGAGGCTATATGTCGACCACGACCCTCAATGGAGGAT
TTCCTCGCTAAAATTTAG

>NER_CYP81E_allele 2
ATGGAAATTTTATTCACTTTTTTATCCATCTTTGCTACTTTTTTCTTCCTCTATAAAATCATCCATTCCCTTAAAACAAACAACAAAA
AACACCCACCAAGCCCACCAAGCCCACCATCATACCCATTATTGGGCCACTTACACCTTCTAAAACCCCCATTCCACCGTACACTC
CAATCTCTAGCCCAACGTTACGGCCCAATCTTCACTCTCAAATTGGGCCTTCAACGGGCCGTGGTAGTTTCCTCCGCATGGG CCGC
CGAAGAATGTTTCGGCCAAAACGACGTCGTTTTTGCAAACAGACCAAAATTCATAATAGGACAACATTTAGGATACAACCACTCT
ATCCTTATCTGGTCCCCATACGGCGACTACTGGCGGAACCTCCGACGTGTCACAACCATTACCATGTTATCCATGAGACGGATCAA
CGAAGCGGGTCCGACCCGGAAACTTGAGATCCGAAACATGATAAGGAACCTTCTAGAAAGTGGGAGTGGGACCCAAAAGGT GAA
TATGAATGAAGTGCTGAGTAAACTAGCAAGAAATTTTGTAATGAGGATAGTAAATGGAAAATCATGGGAAAAAATGATAATAAA
ACCCCCTGAAAATTTAATGACTATTTGTGATTTTTTACCATTTTTAAGATGGGTAGATTTTAAAGGGATAGAAAAGGATATGAAGG
AAAAACAGGTTGAAAGAGATGAGTTTTTACAAAGTTTGGTTGATGAATTTAGAGAAAGTAAGAAGAAAGGTGAGGATTTAGGG A
TGAATAATACTTTGATTGAACAATTGCTGGATTTGCAACAAGCTGAACCTGATCAATACTCTGATGAAACTATCAAAGGGATCATT
TTGGTAATGTTACTAGCGGGATCAGAAACCACAGCACGAACCTTAGAATGGGCATTATCAAATCTCATTAACCACCCAGAAATCC
TCGCGAAAGCAAGAACAGAAATCGACCTCCATGTAGGCAAGGAAAGACTAGTAGATGATTCTGATCTCCCAAAATTAACGTATAC
ACGTTGCATTGTATACGAAACTCTAAGACTATTTCCGGCTGCACCACTTTTAGTACCTCATTTTTCTTCACAAGATTGCACCATAGG
AGGCTATCACGTACCAAAAGGGACAATGCTTTTCGTAAATGCTTGGGCTATACACAGAGACCCGACCCTATGGGACGAGCCTACT
TTGTTTAAGCCCGAGAGGTTCGAAAAGGAGAAAGAAGGGTTTAAGTTTATGCCCTTTGGAATTGGAAGGAGATCTTGTCCAGGGA
ACAATGTGGCTCTTAGGAACGTTTCATTGACATTGGCTACTCTTATTCAGTGCTTTGATTGGGAAGCAGCTGAGAGTGGATCAATT
GATCTAACCGAAAAGCGTGGTGTTGGTGCGATCATTGTCCCTAAGGAAAAGCCATTGGAGGCTATATGTCGACCTCGACCCTCAA
TGGAGGATTTCCTCGCTAAAATTTAG

CLUSTAL alignment

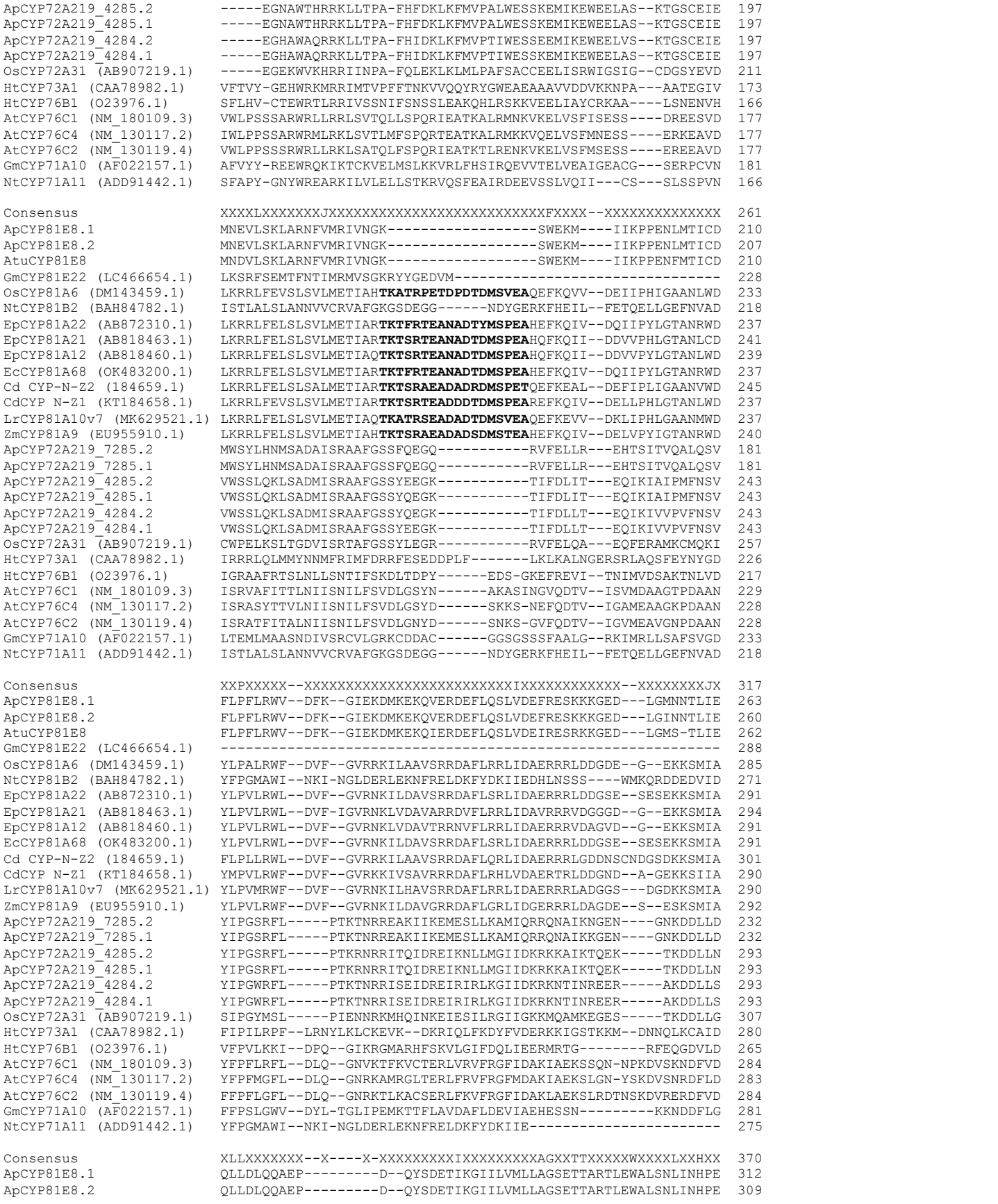
```
CYP81E8 genome  MEILFTFLSIFATFFFLYKIIQSLKPNNKKHPPSPPSPPSPPSYPLLGHLHLLKPPFHRT   60
NES allele1     MEILFTFLSIFATFFFLYKIIHSLKTNNKKHPPSPP---SPPSYPLLGHLHLLKPPFHRT   57
NES allele2     MEILFTFLSIFATFFFLYKIIQSLKTNNKKHPPS------PPSYPLLGHLHLLKPPFHRT   54
NER allele1     MEILFTFLSIFATFFFLYKIIHSLKTNNKKHPPSPP---SPPSYPLLGHLHLLKPPFHRT   57
NER allele2     MEILFTFLSIFATFFFLYKIIQSLKTNNKKHPPS------PPSYPLLGHLHLLKPPFHRT   54

CYP81E8 genome  LQSLAQRYGPIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL  120
NES allele1     LQSLAQRYGPIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL  117
NES allele2     LQSLAQRYGPIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL  114
NER allele1     LQSLAQRYGPIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL  117
NER allele2     LQSLAQRYGPIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL  114

CYP81E8 genome  IWSPYGDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIMNLLESGSGTQKVNMNEVL  180
NES allele1     IWSPYGDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIRNLLESGSGTQKVNMNEVL  177
NES allele2     IWSPYGDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIMNLLESGSGTQKVNMNEVL  174
NER allele1     IWSPYGDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIRNLLESGSGTQKVNMNEVL  177
NER allele2     IWSPYGDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIMNLLESGSGTQKVNMNEVL  174

CYP81E8 genome  SKLARNFVMRIVNGKSWEKMIIKPPENLMTICDFLPFLRWVDFKGIEKDMKEKQVERDEF  240
NES allele1     SKLARNFVMRIVNGKSWEKMIIKPPENLMTICDFLPFLRWVDFKGIEKDMKEKQVERDEF  237
NES allele2     SKLARNFVMRIVNGKSWEKMIIKPPENLMTICDFLPFLRWVDFKGIEKDMKEKQVERDEF  234
NER allele1     SKLARNFVMRIVNGKSWEKMIIKPPENLMTICDFLPFLRWVDFKGIEKDMKEKQVERDEF  237
NER allele2     SKLARNFVMRIVNGKSWEKMIIKPPENLMTICDFLPFLRWVDFKGIEKDMKEKQVERDEF  234

CYP81E8 genome  LQSLVDEFRESKKKGEDLGMNNTLIEQLLDLQQAEPDQYSDETIKGIILVMLLAGSETTA  300
NES allele1     LQSLVDEFRESKKKGEDLGMNNTLIEQLLDLQQAEPDQYSDETIKGIILVMLLAGSETTA  297
NES allele2     LQSLVDEFRESKKKGEDLGINNTLIEQLLDLQQAEPDQYSDETIKGIILVMLLAGSETTA  294
NER allele1     LQSLVDEFRESKKKGEDLGMNNTLIEQLLDLQQAEPDQYSDETIKGIILVMLLAGSETTA  297
NER allele2     LQSLVDEFRESKKKGEDLGINNTLIEQLLDLQQAEPDQYSDETIKGIILVMLLAGSETTA  294
```



Clustal alignment of cytochrome P450 involved in herbicide metabolism. Ap – Amaranthus palmeri, Atu – *Amaranthus tuberculatus*, At – *Arabidopsis thaliana*, Cd – *Cynodon dactylon*, Ec – *Echinochloa crus-galli*, Ep – *Echinochloa phyllogon*, Gm – *Glycine max*, Ht – *Helianthus tuberorus*, Lr – *Lolium rigidum*, Nt – *Nitotiana tabacum*, Os – *Oryza sativa*, Zm – *Zea mays*.

```
Consensus                MXXXXXXLXXXPXEXMXXXXX-XXXXXXXXX---XXXXXXXXXXXXXXXXX---XXXXXX   53
ApCYP81E8.1              ---------------MEILFT---------F---LSIFATFFFLYKIIHS------LKTN   27
ApCYP81E8.2              ---------------MEILFT---------F---LSIFATFFFLYKIIQS------LKTN   27
AtuCYP81E8               ---------------MEILYT---------Y---LAIFSTIFFLYKIIHS------LKPN   27
GmCYP81E22 (LC466654.1)  ---------------MTPFYF-LLF---A-F---ILFLSINFLIQT--------------   23
OsCYP81A6 (DM143459.1)   ---------------MDNAYIIAIL---S-V---AILFLLHYYLLGRGNG----------   28
NtCYP81B2 (BAH84782.1)   -----------------MKFL-LVV---A-S---LFLFVFLILSATKR------------   23
EpCYP81A22 (AB872310.1)  ---------------MDKAYI-AVI---S-I---AFLFLLNYLLGRGGGGR---RNG---   31
EpCYP81A21 (AB818463.1)  ---------------MDKAYV-AVL---T-F---AFLAVLHYLVRRRVGGN---ANGKGR   34
EpCYP81A12 (AB818460.1)  ---------------MDKAYV-AVL---T-F---ALLAVLHYLVR-RVGGN---ANGKGR   33
EcCYP81A68 (OK483200.1)  ---------------MDKAYI-AVI---S-I---AFLFLLNYLLGRGGGGR---RNG---   31
Cd CYP-N-Z2 (184659.1)   ---------------MDKAYVAIVL---S-I---LFLFSIQRFLGHRRRSRSNVDNGKNK   38
CdCYP N-Z1 (KT184658.1)  ---------------MDKAYV-ALL---S-F---ASLFLLHYLVSRRNGTG---KGSKA-   33
LrCYP81A10v7 (MK629521.1) ---------------MDKAYI-AIL---S-C---AFLFLVHYVLGKVSHGR---RGKK--   32
ZmCYP81A9 (EU955910.1)   ---------------MDKAYI-AAL---S-A---AALFLLHYLLGRRAGGE---GKAKAK   34
ApCYP72A219_7285.2       ------------------------------------------------------------
ApCYP72A219_7285.1       ------------------------------------------------------------
ApCYP72A219_4285.2       ------------MEEKSILSI--GISLFCVILFSVSWKIINWVWLKPKKFEKLL-KNQGF   45
ApCYP72A219_4285.1       ------------MEEKSILSI--GISLFCVILFSVSWKIINWVWLKPKKFEKLL-KNQGF   45
ApCYP72A219_4284.2       ------------MEETSILSI--GISLVCVVLVSVSWGILNWVWFKPKRLEKCL-RNQGF   45
ApCYP72A219_4284.1       ------------MEETSILSI--GISLVCVVLLSVSWVILNWVWFKPKRLEKCL-RNQGF   45
OsCYP72A31 (AB907219.1)  -MVFRGWLMWAPASAPVLVVF--G-LLFGLALVWQAGRLLHRLWWRPRRLEKAL-RARGL   55
HtCYP73A1 (CAA78982.1)   ---------------MDLLLIEKTL---V-A---LFAAIIGAILISKL------------   26
HtCYP76B1 (O23976.1)     ---------------MD----------FLII---VSTLLLSYILIWVLG------------   21
AtCYP76C1 (NM_180109.3)  ---------------MDIISGQALLLLFCFI---LSCFLIFTTTRSGRI-----------   31
AtCYP76C4 (NM_130117.2)  ---------------MDIISGQALFLLFCFI---SSCFLISTTARSRRS-----------   31
AtCYP76C2 (NM_130119.4)  ---------------MDIIFEQALFPLFCFV---LSFFIIFFTTTRPRS-----------   31
GmCYP71A10 (AF022157.1)  MALLSSVLKQLPHELSSTHYL-T-----------VFFCIFLILLQLI-------------   35
NtCYP71A11 (ADD91442.1)  -----------------MKFL-LVV---A-S---LFLFVFLILSATKR------------   23

Consensus                XXXXXXLXX-------------XPXXXPXJGXXXXXXX-XXHXXXXXLXX---XXG----   92
```

```
ApCYP81E8.1                NKKHPPSPP-------------SPPSYPLLGHLHLLKP-PFHRTLQSLAQ---RYG----   66
ApCYP81E8.2                NKKHPPS----------------PPSYPLLGHLHLLKP-PFHRTLQSLAQ---RYG----   63
AtuCYP81E8                 NKKLPP-----------------SPPSYPILGHLHLLKP-PFHRTLQSLAQ---RYG----   63
GmCYP81E22 (LC466654.1)    -RRFKNLPP-------------GPFSFPIIGNLHQLKQ-PLHRTFHALSQ---KYG----   61
OsCYP81A6 (DM143459.1)     --GAARLPP-------------GPPAVPILGHLHLVKK-PMHATMSRLAE---RYG----   65
NtCYP81B2 (BAH84782.1)     KSKAKKLPP-------------GPRKLPVIGNLLQIGK-LPHRSLQKLSN---EYG----   62
EpCYP81A22 (AB872310.1)    Q-GARQLPP-------------SPPAVPFLGHLHLVKT-PFHAALVRLAA---RHG----   69
EpCYP81A21 (AB818463.1)    KAPPFRLPP-------------SPPAVPFLGHLHLVRA-PFHAALARLAA---RHG----   73
EpCYP81A12 (AB818460.1)    KAP-SRLPP-------------SPPAVPFLGHLHLVKA-PFHAALARLAA---RHG----   71
EcCYP81A68 (OK483200.1)    K-GARQLPP-------------SPPAVPFLGHLHLVKT-PFHAALVRLAA---RHG----   69
Cd CYP-N-Z2 (184659.1)     SVTHNRLPP-------------GPRAVPVLGHLHLLKK-PIHAALARLAS---QHG----   77
CdCYP N-Z1 (KT184658.1)    ---KGALPP-------------SPPSVPFLGHLHLVKT-PFHAALARLAD---CHG----   69
LrCYP81A10v7 (MK629521.1)  --GAVQLPP-------------SPPAIPFIGHLHLVEK-PIHATMCRLAA---RLG----   69
ZmCYP81A9 (EU955910.1)     G-SRRRLPP-------------SPPAIPFLGHLHLVKA-PFHGALARLAA---RHG----   72
ApCYP72A219_7285.2         -----------------MVKETTSKPMDHFSNDYIP------RVFPFHHQIIKNYG----   33
ApCYP72A219_7285.1         -----------------MVKETTSKPMDHFSNDYIP------RVFPFHHQIIKNYG----   33
ApCYP72A219_4285.2         DGNSYRLLHGDTKDKAKMTMEVTSKPMPNLSNDFLP------HTLAFFHHTISKFG----   95
ApCYP72A219_4285.1         DGNSYRLLHGDTKDKAKMTMEVTSKPMPNLSNDFLP------HTLAFFHHTISKFG----   95
ApCYP72A219_4284.2         NGNSYRLLYGDTNERARMTMESRLNPMPNLSNHFLP------HSLAFYHHIITKYG----   95
ApCYP72A219_4284.1         NGNSYRLLYGDTNERARMTMESRLNPMPNLSNHFLP------HSLAFFHHIITKYG----   95
OsCYP72A31 (AB907219.1)    RGSSYRFLTGDLAEESRRRKEAWARPLPLRCHDIAP------RIKPFLHDTLGEHGKQR-  108
HtCYP73A1 (CAA78982.1)     RGKKFKLPP-------------GPIPVPIFGNWLQVGDDLNHRNLTDLAK---RFG----   66
HtCYP76B1 (O23976.1)       VGKPKNLPP-------------GPTRLPIIGNLHLLGA-LPHQSLAKLAK---IHG----   60
AtCYP76C1 (NM_180109.3)    SRGATALPP-------------GPPRLPIIGNIHLVGK-HPHRSFAELSK---TYG----   70
AtCYP76C4 (NM_130117.2)    SGRAATLPP-------------GPPRLPIIGNIHQVGK-NPHSSFADLAK---IYG----   70
AtCYP76C2 (NM_130119.4)    SRKVVPSPP-------------GPPRLPIIGNIHLVGR-NPHHSFADLSK---TYG----   70
GmCYP71A10 (AF022157.1)    RRNKYNLPP-------------SPPKIPIIGNLHQLGT-LPHRSFHALSH---KYG----   74
NtCYP71A11 (ADD91442.1)    KSKAKKLPP-------------GPRKLPVIGNLLQIGK-LPHRSLQKLSN---EYG----   62

Consensus                  --------XXXXJXXGXXXXVVXXSXXXAXEXXXXXXXXFXXRXXXXXXXXXXXXXXXXX-  143
ApCYP81E8.1                --------PIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL-  117
ApCYP81E8.2                --------PIFTLKLGLQRAVVVSSAWAAEECFGQNDVVFANRPKFIIGQHLGYNHSIL-  114
AtuCYP81E8                 --------PIFSLQLGLQRALVVSSAWAAEECFSQNDVVFANRPKFIVGQHLGYNHSIL-  114
GmCYP81E22 (LC466654.1)    --------PIFSLWFGSRFVVVVSSPLAVQECFTKNDIVLANRPHFLTGKYIGYNNTTV-  112
OsCYP81A6 (DM143459.1)     --------PVFSLRLGSRRAVVVSSPGCARECFTEHDVTFANRPRFESQLLVSFNGAAL-  116
NtCYP81B2 (BAH84782.1)     --------DFIFLQLGSVPTVVVFSAGIAREIFRTQDLVFSGRPALYAGKRFSYNCCNV-  113
EpCYP81A22 (AB872310.1)    --------PVFSLRMGSRRAVVVSSPECARECFTEHDVSFANRPLFPSMRLLAFDGAML-  120
EpCYP81A21 (AB818463.1)    --------PVFTMRMGSRRAVVVSSPEYARECFTEHDVAFANRPLFPTQELVSFSGASL-  124
EpCYP81A12 (AB818460.1)    --------PVFTMRMGSRRAVVVSSPECARECFTEHDVSFANRPLFPTQEIVSFSGASL-  122
EcCYP81A68 (OK483200.1)    --------LVFSLRMGSRRAVVVSSPECARECFTEHDVSFANRPLFPSMRLLAFDGAML-  120
Cd CYP-N-Z2 (184659.1)     --------PLFSLRLGSRPAVVVTSAELARECFTEHDVTFATRPRFASLDLVSFGGTTL-  128
CdCYP N-Z1 (KT184658.1)    --------PVFSLRMGARPAVVVSSPEHAKECFTEHDVAFANRPRFPSQQLASFNGAAL-  120
LrCYP81A10v7 (MK629521.1)  --------PVFSLRLGSRRAVVVSSSECARECFTEHDVTFANRPKFPSQLLASFNGTAL-  120
ZmCYP81A9 (EU955910.1)     --------PVFSMRLGTRRAVVVSSPDCARECFTEHDVNFANRPLFPSMRLASFDGAML-  123
ApCYP72A219_7285.2         --------EKSFVWDGPIPSIMITNPEIIREILMK-IYDFKKIVPTPLMKRLTNGLVRL-   83
ApCYP72A219_7285.1         --------EKSFVWDGPIPSIMITNPEEIRIILMK-IYDFKKIVPTPLMKRLTNGLVRL-   83
ApCYP72A219_4285.2         --------RKCMIWNGPVPFASIAEPDLVREVFMK-INQFQKPKLNPTANKVISGLITL-  145
ApCYP72A219_4285.1         --------RKCMIWNGPVPFASIAEPDLVREVFMK-INQFQNPKLNPIANKVISGLITL-  145
ApCYP72A219_4284.2         --------RRCFIWNGPIPLVNIEEPNLVREIFMK-INDFQKAKSNPIIKKVAPGLLFL-  145
ApCYP72A219_4284.1         --------RRCFIWNGPIPLVNIEEPYLIREIFMK-INDFQKAKSNPIIKKVAPGLLFL-  145
OsCYP72A31 (AB907219.1)    --------QPCITWFGPTPEVNITDPELAKVVLSNKFGHLERVRFKEVSKLLSQGLTYH-  159
HtCYP73A1 (CAA78982.1)     --------EILLLRMGQRNLVVVSSPELAKEVLHTQGVEFGSRTRNVVFDIFTGKGQDM-  117
HtCYP76B1 (O23976.1)       --------PIMSLQLGQITTLVISSATAAEEVLKKQDLAFSTRNVPDAVRAYNHERHSI-  111
AtCYP76C1 (NM_180109.3)    --------PVMSLKLGSLNTVVIASPEAAREVLRTHDQILSARSPTNAVRSINHQDASL-  121
AtCYP76C4 (NM_130117.2)    --------PIMSLKFGCLNSVVITSPEAAREVLRTHDQILSGRKSNDSIRCFGHEEVSV-  121
AtCYP76C2 (NM_130119.4)    --------PIMSLKFGSLNTVVVTSPEAAREVLRTYDQILSSRTPTNSIRSINHDKVSV-  121
GmCYP71A10 (AF022157.1)    --------PLMMLQLGQIPTLVVSSADVAREIIKTHDVVFSNRRQPTAAKIFGYGCKDV-  125
NtCYP71A11 (ADD91442.1)    --------DFIFLQLGSVPTVVVFSAGIAREIFRTQDLVFSGRPALYAGKRFSYNCCNV-  113

Consensus                  XXXXXXGXXWRXXRXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXVX  203
ApCYP81E8.1                IWSPY-GDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIRNLLE---SG-SGTQKVN  172
ApCYP81E8.2                IWSPY-GDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIMNLLE---SG-SGTQKVN  169
AtuCYP81E8                 IWSPY-GDYWRNLRRVTTITMLSMRRINEAGPTRKLEIRNMIMGLLETSKGG-SGTQKVN  172
GmCYP81E22 (LC466654.1)    AVSPY-GDHWRNLRRIMALEVLSTHRINSFLENRRDEIMRLVQKLA---RDSRNGFTKVE  168
OsCYP81A6 (DM143459.1)     ATASY-GAHWRNLRRIVAVQLLSAHRVGLMSGLIAGEVRAMVRRMYRAAAASPAGAARIQ  175
NtCYP81B2 (BAH84782.1)     SFAPY-GNYWREARKILVLELLSTKRVQSFEAIRDEEVSSLVQII---CS---SLSSPVN  166
EpCYP81A22 (AB872310.1)    FVSSY-GPHWRNLRRVAAVQLLSAHRVACMTPAIAAEVRAMVRRMDHAAAAAPGGAARVQ  179
EpCYP81A21 (AB818463.1)    SMASY-GPYWRNLRRVAAVQLLSTHRVACMSPVISAEVRAMVRRMDRAAAAAAGGAAQVQ  183
EpCYP81A12 (AB818460.1)    SMASY-GPYWRNLRRVAAVQLLSTHRVACMSPVISAEVRAMVRRMDRAAAAAAGGAARVQ  181
EcCYP81A68 (OK483200.1)    FVSSY-GPHWRNLRRVAAVQLLSAHRVACMTPAIAAEVRAMVRRMDHAAAAAPGGAARVQ  179
Cd CYP-N-Z2 (184659.1)     PTSRY-GPYWRNLRRVATVHLLSAHRVGCMLPVVSSEVRAMARRVYRAAAAAPRGAARVE  187
CdCYP N-Z1 (KT184658.1)    GSASY-GPYWRNLRRVATVHLLSAHRVACMTGTIAAEVRAMVRRMNRAAQVASGGAARIE  179
LrCYP81A10v7 (MK629521.1)  VTSSY-GPHWRNLRRVATVQLLSAHRVACMSGVIAAEVRAMARRLFHAAEASPDGAARVQ  179
ZmCYP81A9 (EU955910.1)     SVSSY-GPYWRNLRRVAAVQLLSAHRVGCMAPAIEAQVRAMVRRMDRAAAAGGGGVARVQ  182
ApCYP72A219_7285.2         -----EGTKWAKERKLLNPA-FHMEKLKHMIPAFHASCKEMIKEWEKIIS--KTGSGEVE  135
ApCYP72A219_7285.1         -----EGTKWAKERKLLNPA-FHMEKLKHMIPAFHASCKEMIKEWEKIIS--KTGNGEVE  135
```


ApCYP72A219_4285.2 -----EGNAWTHRRKLLTPA-FHFDKLKFMVPALWESSKEMIKEWEELAS--KTGSCEIE 197
ApCYP72A219_4285.1 -----EGNAWTHRRKLLTPA-FHFDKLKFMVPALWESSKEMIKEWEELAS--KTGSCEIE 197
ApCYP72A219_4284.2 -----EGHAWAQRRKLLTPA-FHIDKLKFMVPTIWESSEEMIKEWEELVS--KTGSCEIE 197
ApCYP72A219_4284.1 -----EGHAWAQRRKLLTPA-FHIDKLKFMVPTIWESSKEMIKEWEELAS--KTGSCEIE 197
OsCYP72A31 (AB907219.1) -----EGEKWVKHRRIINPA-FQLEKLKLMLPAFSACCEELISRWIGSIG--CDGSYEVD 211
HtCYP73A1 (CAA78982.1) VFTVY-GEHWRKMRRIMTVPFFTNKVVQQYRYGWEAEAAAVVDDVKKNPA---AATEGIV 173
HtCYP76B1 (O23976.1) SFLHV-CTEWRTLRRIVSSNIFSNSSLEAKQHLRSKKVEELIAYCRKAA----LSNENVH 166
AtCYP76C1 (NM_180109.3) VWLPSSSARWRLLRRLSVTQLLSPQRIEATKALRMNKVKELVSFISESS----DREESVD 177
AtCYP76C4 (NM_130117.2) IWLPPSSARWRMLRKLSVTLMFSPQRTEATKALRMKKVQELVSFMNESS----ERKEAVD 177
AtCYP76C2 (NM_130119.4) VWLPPSSSRWRLLRKLSATQLFSPQRIEATKTLRENKVKELVSFMSESS----EREEAVD 177
GmCYP71A10 (AF022157.1) AFVYY-REEWRQKIKTCKVELMSLKKVRLFHSIRQEVVTELVEAIGEACG---SERPCVN 181
NtCYP71A11 (ADD91442.1) SFAPY-GNYWREARKILVLELLSTKRVQSFEAIRDEEVSSLVQII---CS---SLSSPVN 166
Consensus XXXXLXXXXXXXJXXXXXXXXXXXXXXXXXXXXXXXXXXXFXXXX--XXXXXXXXXXXXXX 261
ApCYP81E8.1 MNEVLSKLARNFVMRIVNGK------------------SWEKM----IIKPPENLMTICD 210
ApCYP81E8.2 MNEVLSKLARNFVMRIVNGK------------------SWEKM----IIKPPENLMTICD 207
AtuCYP81E8 MNDVLSKLARNFVMRIVNGK------------------SWEKM----IIKPPENFMTICD 210
GmCYP81E22 (LC466654.1) LKSRFSEMTFNTIMRMVSGKRYYGEDVM------------------------------ 228
OsCYP81A6 (DM143459.1) LKRRLFEVSLSVLMETIAH**TKATRPETDPDTDMSVEA**QEFKQVV--DEIIPHIGAANLWD 233
NtCYP81B2 (BAH84782.1) ISTLALSLANNVVCRVAFGKGSDEGG------NDYGERKFHEIL--FETQELLGEFNVAD 218
EpCYP81A22 (AB872310.1) LKRRLFELSLSVLMETIAR**TKTFRTEANADTYMSPEA**HEFKQIV--DQIIPYLGTANRWD 237
EpCYP81A21 (AB818463.1) LKRRLFELSLSVLMETIAR**TKTSRTEANADTDMSPEA**HQFKQII--DDVVPHLGTANLCD 241
EpCYP81A12 (AB818460.1) LKRRLFELSLSVLMETIAQ**TKTSRTEANADTDMSPEA**HQFKQII--DDVVPYLGTANLWD 239
EcCYP81A68 (OK483200.1) LKRRLFELSLSVLMETIAR**TKTFRTEANADTDMSPEA**HEFKQIV--DQIIPYLGTANRWD 237
Cd CYP-N-Z2 (184659.1) LKRRLFELSLSALMETIAR**TKTSRAEADADRDMSPET**QEFKEAL--DEFIPLIGAANVWD 245
CdCYP N-Z1 (KT184658.1) LKRRLFEVSLSVLMETIAR**TKTSRTEADDDTDMSPEA**REFKQIV--DELLPHLGTANLWD 237
LrCYP81A10v7 (MK629521.1) LKRRLFELSLSVLMETIAQ**TKATRSEADADTDMSVEA**QEFKEVV--DKLIPHLGAANMWD 237
ZmCYP81A9 (EU955910.1) LKRRLFELSLSVLMETIAH**TKTSRAEADADSDMSTEA**HEFKQIV--DELVPYIGTANRWD 240
ApCYP72A219_7285.2 MWSYLHNMSADAISRAAFGSSFQEGQ-----------RVFELLR---EHTSITVQALQSV 181
ApCYP72A219_7285.1 MWSYLHNMSADAISRAAFGSSFQEGQ-----------RVFELLR---EHTSITVQALQSV 181
ApCYP72A219_4285.2 VWSSLQKLSADMISRAAFGSSYEEGK-----------TIFDLIT---EQIKIAIPMFNSV 243
ApCYP72A219_4285.1 VWSSLQKLSADMISRAAFGSSYQEGK-----------TIFDLIT---EQIKIAIPMFNSV 243
ApCYP72A219_4284.2 VWSSLQKLSADMISRAAFGSSYQEGK-----------TIFDLLT---EQIKIVVPVFNSV 243
ApCYP72A219_4284.1 VWSSLQKLSADMISRAAFGSSYEEGK-----------TIFDLLT---EQIKIVVPVFNSV 243
OsCYP72A31 (AB907219.1) CWPELKSLTGDVISRTAFGSSYLEGR-----------RVFELQA---EQFERAMKCMQKI 257
HtCYP73A1 (CAA78982.1) IRRRLQLMMYNNMFRIMFDRRFESEDDPLF-------LKLKALNGERSRLAQSFEYNYGD 226
HtCYP76B1 (O23976.1) IGRAAFRTSLNLLSNTIFSKDLTDPY------EDS-GKEFREVI--TNIMVDSAKTNLVD 217
AtCYP76C1 (NM_180109.3) ISRVAFITTLNIISNILFSVDLGSYN------AKASINGVQDTV--ISVMDAAGTPDAAN 229
AtCYP76C4 (NM_130117.2) ISRASYTTVLNIISNILFSVDLGSYD------SKKS-NEFQDTV--IGAMEAAGKPDAAN 228
AtCYP76C2 (NM_130119.4) ISRATFITALNIISNILFSVDLGNYD------SNKS-GVFQDTV--IGVMEAVGNPDAAN 228
GmCYP71A10 (AF022157.1) LTEMLMAASNDIVSRCVLGRKCDDAC------GGSGSSSFAALG--RKIMRLLSAFSVGD 233
NtCYP71A11 (ADD91442.1) ISTLALSLANNVVCRVAFGKGSDEGG------NDYGERKFHEIL--FETQELLGEFNVAD 218
Consensus XXPXXXXX--XXXXXXXXXXXXXXXXXXXXXXXXIXXXXXXXXXXXX--XXXXXXXJX 317
ApCYP81E8.1 FLPFLRWV--DFK--GIEKDMKEKQVERDEFLQSLVDEFRESKKKGED---LGMNNTLIE 263
ApCYP81E8.2 FLPFLRWV--DFK--GIEKDMKEKQVERDEFLQSLVDEFRESKKKGED---LGINNTLIE 260
AtuCYP81E8 FLPFLRWV--DFK--GIEKDMKEKQIERDEFLQSLVDEIRESRKKGED---LGMS-TLIE 262
GmCYP81E22 (LC466654.1) ---------- 288
OsCYP81A6 (DM143459.1) YLPALRWF--DVF--GVRRKILAAVSRRDAFLRRLIDAERRRLDDGDE--G--EKKSMIA 285
NtCYP81B2 (BAH84782.1) YFPGMAWI--NKI-NGLDERLEKNFRELDKFYDKIIEDHLNSSS----WMKQRDDEDVID 271
EpCYP81A22 (AB872310.1) YLPVLRWL--DVF--GVRNKILDAVSRRDAFLSRLIDAERRRLDDGSE--SESEKKSMIA 291
EpCYP81A21 (AB818463.1) YLPVLRWL--DVF-IGVRNKLVDAVARRDVFLRRLIDAVRRRVDGGGD--G--EKKSMIA 294
EpCYP81A12 (AB818460.1) YLPVLRWL--DVF--GVRNKLVDAVTRRNVFLRRLIDAERRRVDAGVD--G--EKKSMIA 291
EcCYP81A68 (OK483200.1) YLPVLRWL--DVF--GVRNKILDAVSRRDAFLSRLIDAERRRLDDGSE--SESEKKSMIA 291
Cd CYP-N-Z2 (184659.1) FLPLLRWL--DVF--GVRRKILAAVSRRDAFLQRLIDAERRRLGDDNSCNDGSDKKSMIA 301
CdCYP N-Z1 (KT184658.1) YMPVLRWF--DVF--GVRKKIVSAVRRRDAFLRHLVDAERTRLDDGND--A-GEKKSIIA 290
LrCYP81A10v7 (MK629521.1) YLPVMRWF--DVF--GVRNKILHAVSRRDAFLRRLIDAERRRLADGGS---DGDKKSMIA 290
ZmCYP81A9 (EU955910.1) YLPVLRWF--DVF--GVRNKILDAVGRRDAFLGRLIDGERRRLDAGDE--S--ESKSMIA 292
ApCYP72A219_7285.2 YIPGSRFL-----PTKTNRREAKIIKEMESLLKAMIQRRQNAIKNGEN----GNKDDLLD 232
ApCYP72A219_7285.1 YIPGSRFL-----PTKTNRREAKIIKEMESLLKAMIQRRQNAIKKGEN----GNKDDLLD 232
ApCYP72A219_4285.2 YIPGSRFL-----PTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEK-----TKDDLLN 293
ApCYP72A219_4285.1 YIPGSRFL-----PTKRNRRITQIDREIKNLLMGIIDKRKKAIKTQEK-----TKDDLLN 293
ApCYP72A219_4284.2 YIPGWRFL-----PTKTNRRISEIDREIRIRLKGIIDKRKNTINREER-----AKDDLLS 293
ApCYP72A219_4284.1 YIPGWRFL-----PTKTNRRISEIDREIRIRLKGIIDKRKNTINREER-----AKDDLLS 293
OsCYP72A31 (AB907219.1) SIPGYMSL-----PIENNRKMHQINKEIESILRGIIGKKMQAMKEGES-----TKDDLLG 307
HtCYP73A1 (CAA78982.1) FIPILRPF--LRNYLKLCKEVK--DKRIQLFKDYFVDERKKIGSTKKM--DNNQLKCAID 280
HtCYP76B1 (O23976.1) VFPVLKKI--DPQ--GIKRGMARHFSKVLGIFDQLIEERMRTG--------RFEQGDVLD 265
AtCYP76C1 (NM_180109.3) YFPFLRFL--DLQ--GNVKTFKVCTERLVRVFRGFIDAKIAEKSSQN-NPKDVSKNDFVD 284
AtCYP76C4 (NM_130117.2) YFPFMGFL--DLQ--GNRKAMRGLTERLFRVFRGFMDAKIAEKSLGN-YSKDVSNRDFLD 283
AtCYP76C2 (NM_130119.4) FFPFLGFL--DLQ--GNRKTLKACSERLFKVFRGFIDAKLAEKSLRDTNSKDVRERDFVD 284
GmCYP71A10 (AF022157.1) FFPSLGWV--DYL-TGLIPEMKTTFLAVDAFLDEVIAEHESSN---------KKNDDFLG 281
NtCYP71A11 (ADD91442.1) YFPGMAWI--NKI-NGLDERLEKNFRELDKFYDKIIE---------- 275
Consensus XLLXXXXXXX--X----X-XXXXXXXXXIXXXXXXXXXAGXXTTXXXXXWXXXXLXXHXX 370
ApCYP81E8.1 QLLDLQQAEP---------D--QYSDETIKGIILVMLLAGSETTARTLEWALSNLINHPE 312
ApCYP81E8.2 QLLDLQQAEP---------D--QYSDETIKGIILVMLLAGSETTARTLEWALSNLINHPE 309

```
AtuCYP81E8                QLLDLQEAEP---------D--QYTDETIKGIILVMLLAGSETTARTLEWALSNLINHPE  311
GmCYP81E22 (LC466654.1)   ------------------------------------------------------------  348
OsCYP81A6 (DM143459.1)    VLLTLQKTEP---------E--VYTDNMITALTANLFGAGTETTSTTSEWAMSLLLNHPD  334
NtCYP81B2 (BAH84782.1)    VLLRIQKDPN---------QEIPLKDDHIKGLLADIFIAGTDTSSTTIEWAMSELIKNPR  322
EpCYP81A22 (AB872310.1)   VLLTLQKSEP---------E--VYTDTMIMALCTNLFGAGTETTSTTTEWAMSLLLNHPE  340
EpCYP81A21 (AB818463.1)   VLLSLQKSEP---------E--VYTDTMIMSLCANLFGAGTETTSTTTEWAMALLLNHPE  343
EpCYP81A12 (AB818460.1)   VLLSLQKSEP---------E--VYTDTTIMSLCANLFGAGTETTSTTTEWAMALLLNHPE  340
EcCYP81A68 (OK483200.1)   VLLTLQKSEP---------E--VYTDTMIMALCTNLFGAGTETTSTTTEWAMSLLLNHPE  340
Cd CYP-N-Z2 (184659.1)    VLLNLQKTEP---------E--VYTDATIMALCTSMFTGGAETTATTTEWAMSLLLNHPD  350
CdCYP N-Z1 (KT184658.1)   MLLTLQKSEP---------D--VYSDTMIMALCGNLFGAGTETTSTTTEWAMSLLLNHPE  339
LrCYP81A10v7 (MK629521.1) VLLTLQKTEP---------K--VYTDTMITALCANLFGAGTETTSTTTEWAMSLLLNHPA  339
ZmCYP81A9 (EU955910.1)    VLLTLQKSEP---------E--VYTDTVITALCANLFGAGTETTSTTTEWAMSLLLNHRE  341
ApCYP72A219_7285.2        LLLTSSIEDNHVKDGNTK-EYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQN  291
ApCYP72A219_7285.1        LLLTSSIEDNHVKDGNTK-EYVKLGMKEIIEECKMFYIAGQETTSSLLTWTLILLSKHQN  291
ApCYP72A219_4285.2        TLLESSYEQEENGHGNKKNN-INLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQN  352
ApCYP72A219_4285.1        TLLESSYEQEENGHGNKKND-INLNIEEILGECKLFYLAGQDTTASLLVWTLIMLAKHQN  352
ApCYP72A219_4284.2        TLIQDNF-----GHANKKQDINKLNIEEIVNECKLFYLAGQDTTSSLLLWTLIMLGKHQI  348
ApCYP72A219_4284.1        TLIQDNF-----GHANKKQDINKLSIEEVVNECKLFYLAGQDTTSSLLLWTLIMLGKHQI  348
OsCYP72A31 (AB907219.1)   ILLESNTKHMEEN----GQSSQGLTMKDIVEECKLFYFAGAETTSVLLTWAMLLLSMHPE  363
HtCYP73A1 (CAA78982.1)    HILEAK--EK---------G--EINEDNVLYIVENINVAAIETTLWSIEWGIAELVNHPE  327
HtCYP76B1 (O23976.1)      VCLKMMQDNP---------N--EFNHTNIKALFLDLFVAGTDTTSITIEWAMTELLRKPH  314
AtCYP76C1 (NM_180109.3)   NLLDY-KGDE---------S--ELSISDIEHLLLVSLTLLQII----MIYKIME------  328
AtCYP76C4 (NM_130117.2)   SLLILNEGDE---------A--ELDNNDIEHLLLDMFTAGTDTSSSTLEWAMAELLRNPK  332
AtCYP76C2 (NM_130119.4)   VLLDLTEGDE---------A--ELNTNDIVHLLLDLFGAGTDTNSSTVEWAMAELLRNPE  333
GmCYP71A10 (AF022157.1)   ILLQLQECGR---------LDFQLDRDNLKAILVDMIIGGSDTTSTTLEWTFAEFLRNPN  332
NtCYP71A11 (ADD91442.1)   ------------------------------------------------------------  335

Consensus                 XXXXAXXEIXXXXGXXXX----------XXXXXXXXLXXXXXIXXEXLRLXPXXXXXXXX  420
ApCYP81E8.1               ILAKARTEIDLHVGKERL----------VDDSDLPKLTYTRCIVYETLRLFPAAPLLVPH  362
ApCYP81E8.2               ILVKARTEIDLHVGKERL----------VDDSDLPKLTYTRCIVYETLRLFPAAPLLVPH  359
AtuCYP81E8                ILAKARTEIDLHVGKERL----------VDDSDLPKLTYTRCIVYETLRLFPAAPLLVPH  361
GmCYP81E22 (LC466654.1)   ------------------------------------------------------------  408
OsCYP81A6 (DM143459.1)    TLKKAQAEIDASVGNSRL----------ITADDVTRLGYLQCIVRETLRLYPAAPMLLPH  384
NtCYP81B2 (BAH84782.1)    VLRKAQEEVREVAKGKQK----------VQESDLCKLEYLKLVIKETLRLHPPAPLLVPR  372
EpCYP81A22 (AB872310.1)   KLKKAQAEIDAAVGTSRL----------VTADDVPRLAYLQSIIHETLRLYPAAPLLLPH  390
EpCYP81A21 (AB818463.1)   KLKKAQAEIDAAVGTSRL----------ITPDDMPHLGYLHSIIHETLRLYPAAPLLLPH  393
EpCYP81A12 (AB818460.1)   KLKRAQAEIDAAVGTSRL----------ITPDVMPRLGYLQSIIHETLRLYPAAPLLLPH  390
EcCYP81A68 (OK483200.1)   KLKKAQAEIDAAVGTSRL----------VTADDVPHLAYLQSIIHETLRLYPAAPLLLPH  390
Cd CYP-N-Z2 (184659.1)    VLKKAQAEMDVSVGTSRL----------VTAADVPHLGYLQCIISETLRLYPAVPTLVPH  400
CdCYP N-Z1 (KT184658.1)   KLRKAQAEIDAVVGTSRL----------LTADDMPRLTYLRCIIDETMRLYPAAPLLLPH  389
LrCYP81A10v7 (MK629521.1) ALKKAQAEIDASVGTSRL----------VSVDDVPSLAYLQCIVSETLRLYPAAPLLLPH  389
ZmCYP81A9 (EU955910.1)    ALKKAQAEIDAAVGTSRL----------VTADDVPHLTYLQCIVDETLRLHPAAPLLLPH  391
ApCYP72A219_7285.2        WQHQAREEILATFGDSIP-----------NDEGLNQLKKVSMILQEVLRLYPPVVT-LGR  339
ApCYP72A219_7285.1        WQHQAREEILATFGDTIP-----------NDEGLNQLKKVSMILQEVLRLYPPVVT-LGR  339
ApCYP72A219_4285.2        WQQLARQEIFNAFGDKIP-----------QFHELNHLKIVNMILQEVLRLYPPVTE-TNR  400
ApCYP72A219_4285.1        WQQLARQEIFNAFGDKTP-----------QFHELNHLKIVNMILQEVLRLYPPVTE-TNR  400
ApCYP72A219_4284.2        WQQRAREEIINTFGDAKP-----------QFHELNSLKIVNMILQEVLRLYPPVLE-LNR  396
ApCYP72A219_4284.1        WQQRAREEIINAFGDAKP-----------QFHELNSLKIVNMILQEVLRLYPPILE-LNR  396
OsCYP72A31 (AB907219.1)   WQDRAREEILGLFRKNKP-----------DYEGLSRLKIVTMILYEVLRLYPPFIE-IGR  411
HtCYP73A1 (CAA78982.1)    IQAKLRHELDTKLGPGVQ----------ITEPDVQNLPYLQAVVKETLRLRMAIPLLVPH  377
HtCYP76B1 (O23976.1)      IMSKAKEELEKVIGKGSI----------VKEDDVLRLPYLSCIVKEVLRLHPPSPLLLPR  364
AtCYP76C1 (NM_180109.3)   ------------------------------------------------------------  388
AtCYP76C4 (NM_130117.2)   TMVKAQAEMDRVLGQNSV----------VQESDISGLPYLQAVVKETFRLHPAAPLLVPR  382
AtCYP76C2 (NM_130119.4)   TMVKAQAEIDCVIGQKGV----------VEESDISALPYLQAVVKETFRLHPAAPLLVPR  383
GmCYP71A10 (AF022157.1)   TMKKAQEEVRRVVGINSKA--------VLDENCVNQMNYLKCVVKETLRLHPPLPLLIAR  384
NtCYP71A11 (ADD91442.1)   ------------------------------------------------------------  395

Consensus                 XXXXDXXXGXX-XXPXGXXXXXXXXXIXXDPXXWXXXXXXFXPERFXXXX---------X  470
ApCYP81E8.1               FSSQDCTIGGY-HVPKGTMLFVNAWAIHRDPTLWD-EPTLFKPERFEKE-----------  409
ApCYP81E8.2               FSSQDCTIGGY-HVPKGTMLFVNAWAIHRDPTLWD-EPTVFKPERFEKE-----------  406
AtuCYP81E8                FSSQDCTIGGY-HVPKGTMLFVNAWAIHRDPTLWD-EPTVFKPERFEKE-----------  408
GmCYP81E22 (LC466654.1)   ------------------------------------------------------------  468
OsCYP81A6 (DM143459.1)    ESSADCKVGGY-NIPRGSMLLINAYAIHRDPAVWE-EPEKFMPERFEDGG----------  432
NtCYP81B2 (BAH84782.1)    VTTASCKIMEY-EIPADTRVLINSTAIGTDPKYWE-NPLTFLPERFLDKE---------I  421
EpCYP81A22 (AB872310.1)   EAAADTRVGGY-DVPRGTMLLVNVYAIHRDPAAWE-DPSAFRPERFEN-G---------K  438
EpCYP81A21 (AB818463.1)   LSSADCNVGGY-DVPRGTMLLVNVYAIHRDPAVWE-DPDAFRPERFEDGK---------A  442
EpCYP81A12 (AB818460.1)   LSSADCKVGGY-DVPRGTMLLVNVYAIHRDPAVWE-DPAAFRPERFEDGK---------A  439
EcCYP81A68 (OK483200.1)   LSSADCKVGGY-DVPRGTMLLVNVYAIHRDPAVWE-DPAAFRPERFEN-G---------K  438
Cd CYP-N-Z2 (184659.1)    ESTADCVIGGHHHVPAGTMLLVNGYAIHRDPATWP-DPAAFRPERFEDG----------  448
CdCYP N-Z1 (KT184658.1)   ESSTHCKVGGY-DVPAGTMLLVNVYAIHRDPAVWD-GPTEFVPERFEDG----------  436
LrCYP81A10v7 (MK629521.1) ESSADCKVGGY-NVPADTMLIVNAYAIHRDPAAWE-HPLEFRPDRFEDG----------  436
ZmCYP81A9 (EU955910.1)    ESAADCTVGGY-DVPRGTMLLVNVHAVHRDPAVWE-DPDRFVPERFEGAG---------G  440
ApCYP72A219_7285.2        KVSHDIKVGDM-DLPAGVQVKIPTIFIHHSEKLWGKDAKEFNPNRFSQGL---------L  389
ApCYP72A219_7285.1        KVSHDIKVGDM-DLPAGVQVKIPTIFIHHSEKLWGKDAKEFNPNRFSQGF---------L  389
ApCYP72A219_4285.2        EVIQDIKLGNM-ILPAGVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGI---------S  450
ApCYP72A219_4285.1        EVIQDIKLGNM-ILPAGVVVNIPILLMHHDEKIWGNDAKEFNPERFSEGI---------S  450
```

```
ApCYP72A219_4284.2        EVTQDIKIGDV-ILPAGVLLNIPILLLQQDEKIWGKDAKEFNPERFSEGI---------S  446
ApCYP72A219_4284.1        EVTQDIKIGDV-ILPAGVLLNIPILLLQQDEKIWGKDAKEFNPERFSEGI---------S  446
OsCYP72A31 (AB907219.1)   KTYKEMEIGGV-TYPAGVSIKIPVLFIHHDPDTWGSDVHEFKPERFSEGI---------S  461
HtCYP73A1 (CAA78982.1)    MNLHDAKLGGF-DIPAESKILVNAWWLANNPDQWK-KPEEFRPERFLEEEA-------KV  428
HtCYP76B1 (O23976.1)      KVVTQVELSGY-TIPAGTLVFVNAWAIGRDPTVWD-DSLEFKPQRFLESR---------L  413
AtCYP76C1 (NM_180109.3)   ------------------------------------------------------------  448
AtCYP76C4 (NM_130117.2)   KAESDVEVLGF-MVPKDTQVLVNVWAIGRDPSVWE-NPSQFEPERFMGKD---------I  431
AtCYP76C2 (NM_130119.4)   KAESDVEVLGF-MVPKDTQVFVNVWAIGRDPNVWE-NSSRFKPERFLGKD---------I  432
GmCYP71A10 (AF022157.1)   ETSSSVKLRGY-DIPAKTMVFINAWAIQRDPELWD-DPEEFIPERFETSQ---------V  433
NtCYP71A11 (ADD91442.1)   ------------------------------------------------------------  455

Consensus                 XXXXXXXXXXXFGXGXRXCXGXXXAXXXXXXLXXJJXXFXXXXXXXXXXXXXXXXXXX  530
ApCYP81E8.1               -K--EGFKFMPFGIGRRSCPGNNVALRNVSLTLATLIQCFDWEAAESG-----SIDLTEK  461
ApCYP81E8.2               -K--EGFKFMPFGIGRRSCPGNNVALRNVSLTLATLIQCFDWEAAESG-----SIDLTEK  458
AtuCYP81E8                -K--EGFKFMPFGIGRRSCPGNNVALRNVSLTLATLIQCFDWEASESG-----SIDLTEK  460
GmCYP81E22 (LC466654.1)   ------------------------------------------------------------  528
OsCYP81A6 (DM143459.1)    -C--DGNLLMPFGMGRRRCPGETLALRTVGLVLGTLIQCFDWERVDGV-----EVDMTEG  484
NtCYP81B2 (BAH84782.1)    DYRGKNFELLPFGAGRRGCPGINFSIPLVELALANLLFHYNWSLPEG--MLPKDVDMEEA  479
EpCYP81A22 (AB872310.1)   AA--EGRLLMPFGMGRRKCPGETLALRTVGLVLGALVQCFDWDRVDGV-----EV--TEG  489
EpCYP81A21 (AB818463.1)   EG--GGRLLMPFGMGRRKCPGETLALRTVGLVLGALIQCFDWDTVDGA-----KVDLTEG  495
EpCYP81A12 (AB818460.1)   E----GRLLMPFGMGRRKCPGEALALRTVGLVLGTLIQCFDWDTVDGA-----KVDLTEG  490
EcCYP81A68 (OK483200.1)   AA--EGRLLMPFGMGRRKCPGETLALRTVGLVLGALVQCFDWDRVDGV-----EVDMTEG  491
Cd CYP-N-Z2 (184659.1)    KA--EGKFIITFGMGRRKCPGETLALRTIGLVLGMLIQCFDWDTADGG-----KVDMTEG  501
CdCYP N-Z1 (KT184658.1)   KA--EGRLLMPFGMGRRKCPGETLALRTIGLVLGTLIQCFDWDRVDGL-----EVDMTES  489
LrCYP81A10v7 (MK629521.1) KA--EGLFMIPFGMGRRRCPGETLALRTIGMVLATLVQCFDWEPVDGV-----KVDMTEG  489
ZmCYP81A9 (EU955910.1)    KA--EGRLLMPFGMGRRKCPGETLALRTVGLVLATLLQCFDWDTVDGA-----QVDMKAS  493
ApCYP72A219_7285.2        KATGGNMCFFAFGWGPRVCIGSNFALLEAKMALALILPHFSFELSPSYAHAPT-IGQG--  446
ApCYP72A219_7285.1        KATGGNMCFFAFGWGPRICIGSNFALLEAKMALALILPHFSFELSPSYAHAPT-IGQG--  446
ApCYP72A219_4285.2        KASKGNMSLFAFGWGPRICIGSNFAMIEAKLTLVCILQRFVFELSPSYLHAPS--AIG--  506
ApCYP72A219_4285.1        KASKGNMSLFAFGWGPRICIGSNFAMIEAKLTLVCILQRFVFELSPSYLHAPS--AIG--  506
ApCYP72A219_4284.2        KASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLVFILQRFVFELSPSYRHAPL--NVG--  502
ApCYP72A219_4284.1        KASKGNMSFFAFGWGPRICLGSNFAMIEAKLTLVFILQRFVFELSPSYRHAPL--NVG--  502
OsCYP72A31 (AB907219.1)   KASKDPGAFLPFGWGPRICIGQNFALLEAKMALCLILQRLEFELAPSYTHAPH--TMV--  517
HtCYP73A1 (CAA78982.1)    EANGNDFRYLPFGVGRRSCPGIILALPILGITIGRLVQNFELLPPPGQ----SKIDTDEK  484
HtCYP76B1 (O23976.1)      DVRGHDFDLIPFGAGRRICPGIPLATRMVPIMLGSLLNNFDWKIDTKV--PYDVLDMTEK  471
AtCYP76C1 (NM_180109.3)   ------------------------------------------------------------  508
AtCYP76C4 (NM_130117.2)   DVKGRDYELTPFGGGRRICPGLPLAVKTVSLMLASLLYSFDWKLPNGV--VSEDLDMDET  489
AtCYP76C2 (NM_130119.4)   DLRGRDYELTPFGAGRRICPGLPLAVKTVPLMLASLLYSFDWKLPNGV--GSEDLDMDET  490
GmCYP71A10 (AF022157.1)   DLNGQDFQLIPFGIGRRGCPAMSFGLASTEYVLANLLYWFNWNMSESGRILMHNIDMSET  493
NtCYP71A11 (ADD91442.1)   ------------------------------------------------------------  515

Consensus                 X---XJXXXXXXXXXXXXXXXXXXVLXXLXE--                            560
ApCYP81E8.1               RGVGAIIVPKEKPLEAICRPRPSMEDFLAKI----                           492
ApCYP81E8.2               RGVGAIIVPKEKPLEAICRPRPSMEDFLAKI----                           489
AtuCYP81E8                RGAGAIIVPKEKPLEAICRPRPSMEDFLVKL----                           491
GmCYP81E22 (LC466654.1)   ----------------------------------                           196
OsCYP81A6 (DM143459.1)    G---GLTIPKVVPLEAMCRPRDAMGGVLRELV---                           513
NtCYP81B2 (BAH84782.1)    L---GITMHKKSPLCLVASHYNLL-----------                           500
EpCYP81A22 (AB872310.1)   G---GLTMPRAVPLEAMCRPRAAMRRVLQEL----                           517
EpCYP81A21 (AB818463.1)   G---GLTIPMAVPLEAMCRPRAALHDVLQEL----                           523
EpCYP81A12 (AB818460.1)   G---GLTIPMAVPLEAMCRPRAALHDVLQEL----                           518
EcCYP81A68 (OK483200.1)   G---GLTMPRAVPLEAMCRPRAAMRRVLQEL----                           519
Cd CYP-N-Z2 (184659.1)    V---GITLPRAVPLEAMCRPRQTMVDVLKGLLE--                           531
CdCYP N-Z1 (KT184658.1)   G---GLTIPRAVPLEAMCRPRATMREVLQEL----                           517
LrCYP81A10v7 (MK629521.1) G---GFTIPKAVPLEAVCRPRAVMRDVLQNL----                           517
ZmCYP81A9 (EU955910.1)    G---GLTMPRAVPLEAMCRPRTAMRGVLKRL----                           521
ApCYP72A219_7285.2        ----TL--RPQFGAKLILHRINRI-----------                           464
ApCYP72A219_7285.1        ----TL--RPQFGAKLILHRINRI-----------                           464
ApCYP72A219_4285.2        ----SL--RPQFGAPIIFRPRILGS----------                           525
ApCYP72A219_4285.1        ----SL--RPQFGAPIIFRPRILGS----------                           525
ApCYP72A219_4284.2        ----TL--RPQFGAPIIFRRR--------------                           517
ApCYP72A219_4284.1        ----TL--RPQFGAPIIFRRR--------------                           517
OsCYP72A31 (AB907219.1)   ----TL--HPMHGAQIKVRAI--------------                           532
HtCYP73A1 (CAA78982.1)    GGQFSLHILKHSTIVAKPRSF--------------                           505
HtCYP76B1 (O23976.1)      N---GTTISKAKPLCVVPIPLN-------------                           490
AtCYP76C1 (NM_180109.3)   ----------------------------------                           322
AtCYP76C4 (NM_130117.2)   F---GITLHRTNTLYAIPVKKQTIN----------                           511
AtCYP76C2 (NM_130119.4)   F---GLTLHKTNPLHAVPVKKRGRN----------                           512
GmCYP71A10 (AF022157.1)   N---GLTVSKKVPLHLEPEPYKT-----------                           513
NtCYP71A11 (ADD91442.1)   ----------------------------------                           252
```

Note: Bold sequence is found in cytochrome P450 *CYP81A* family in grasses and not in dicot. This sequence was inserted in the *CYP81E8v.18aa* to test its function in heterologous expression of this gene in yeast.

REFERENCES

1. Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE and Westra P, Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences* **107**: 1029-1034 (2010).
2. Liu D, Shi L, Han C, Yu J, Li D and Zhang Y, Validation of reference genes for gene expression studies in virus-infected Nicotiana benthamiana using quantitative real-time PCR. *PLoS One* **7**: e46451 (2012).
3. Bailey TL and Elkan C, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36 (1994).
4. Montgomery JS, Giacomini D, Waithaka B, Lanz C, Murphy BP, Campe R, Lerchl J, Landes A, Gatzmann F, Janssen A, Antonise R, Patterson E, Weigel D and Tranel PJ, Draft genomes of *Amaranthus tuberculatus*, *Amaranthus hybridus*, and *Amaranthus palmeri*. *Genome Biol Evol* **12**: 1988-1993 (2020).

6. APPENDIX B

Supporting Information for CHAPTER III - UNRAVELING THE ROLE OF P450

REDUCTASE IN HERBICIDE METABOLIC RESISTANCE MECHANISM

Supporting Information Table 4-1**.** Genotype analysis of F2 segregating populations derived from crosses between *atr1* and *atr2* Arabidopsis mutants.

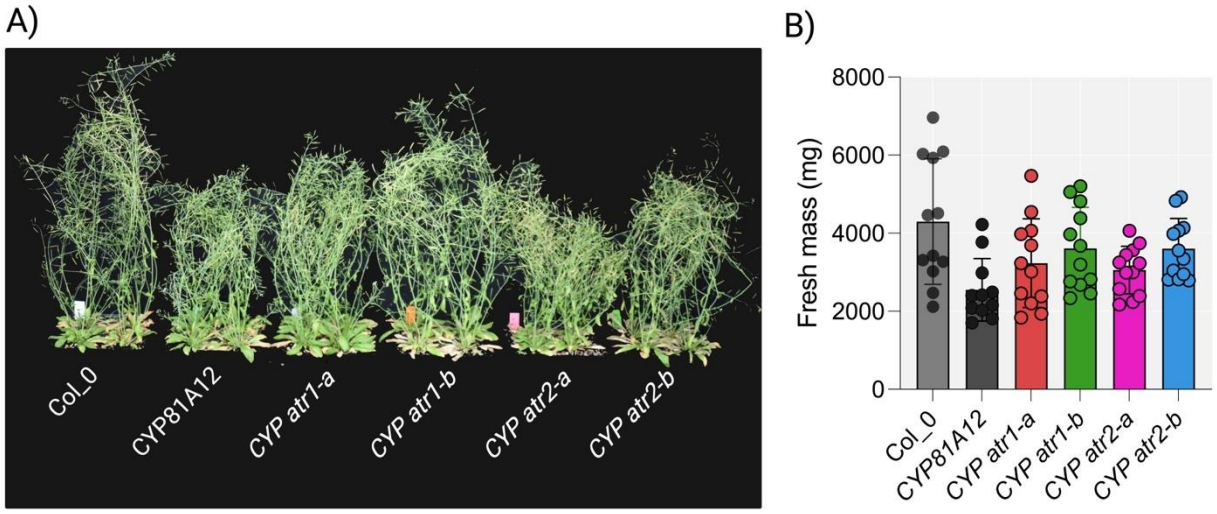
Crosses	F1 plant	ATR1	HZ	HM-WT	HZ	HZ	HM-M	HM-WT	HM-WT	HM-M	HM-M	Total
		ATR2	HZ	HZ	HM-WT	HM-M	HZ	HM-WT	HM-M	HM-WT	HM-M	
atr1-a × *atr2-a*	F1 plant I	Plants	7	2	2	3	0	0	9	14	0	37
		(%)	18.9	5.4	5.4	8.1	0.0	0.0	24.3	37.8	0.0	100
atr1-a × *atr2-a*	F1 plant II	Plants	16	1	2	4	1	0	8	3	0	35
		(%)	45.7	2.9	5.7	11.4	2.9	0.0	22.9	8.6	0.0	100
atr1-a × *atr2-a*	F1 plant III	Plants	27	1	2	3	0	0	12	11	0	56
		(%)	48.2	1.8	3.6	5.4	0.0	0.0	21.4	19.6	0.0	100
atr1-a × *atr2-a*	F1 plant IV	Plants	16	0	0	3	1	0	7	9	0	36
		(%)	44.4	0.0	0.0	8.3	2.8	0.0	19.4	25.0	0.0	100
atr1-b × *atr2-a*	F1 plant I	Plants	20	1	2	1	1	0	16	6	0	47
		(%)	42.6	2.1	4.3	2.1	2.1	0.0	34.0	12.8	0.0	100
atr1-b × *atr2-a*	F1 plant II	Plants	42	2	2	5	2	0	19	14	0	86
		(%)	48.8	2.3	2.3	5.8	2.3	0.0	22.1	16.3	0.0	100
atr1-b × *atr2-b*	F1 plant I	Plants	64	1	3	1	1	0	27	22	0	119
		(%)	53.8	0.8	2.5	0.8	0.8	0.0	22.7	18.5	0.0	100
Total		Plants	192	8	13	20	6	0	98	79	0	416
		(%)	46.2	1.9	3.1	4.8	1.4	0.0	23.6	19.0	0.0	100
Expected		(%)	25	12.5	12.5	12.5	12.5	6.25	6.25	6.25	6.25	100
χ^2 test		$\chi^2 = 560.75$			df = 8				p-value = <0.0001			

HM-WT, homozygous wildtype, HM-MT, homozygous mutant, HZ – heterozygous. *ATR1* – Arabidopsis P450 reductase 1, *ATR2* - Arabidopsis P450 reductase 2.

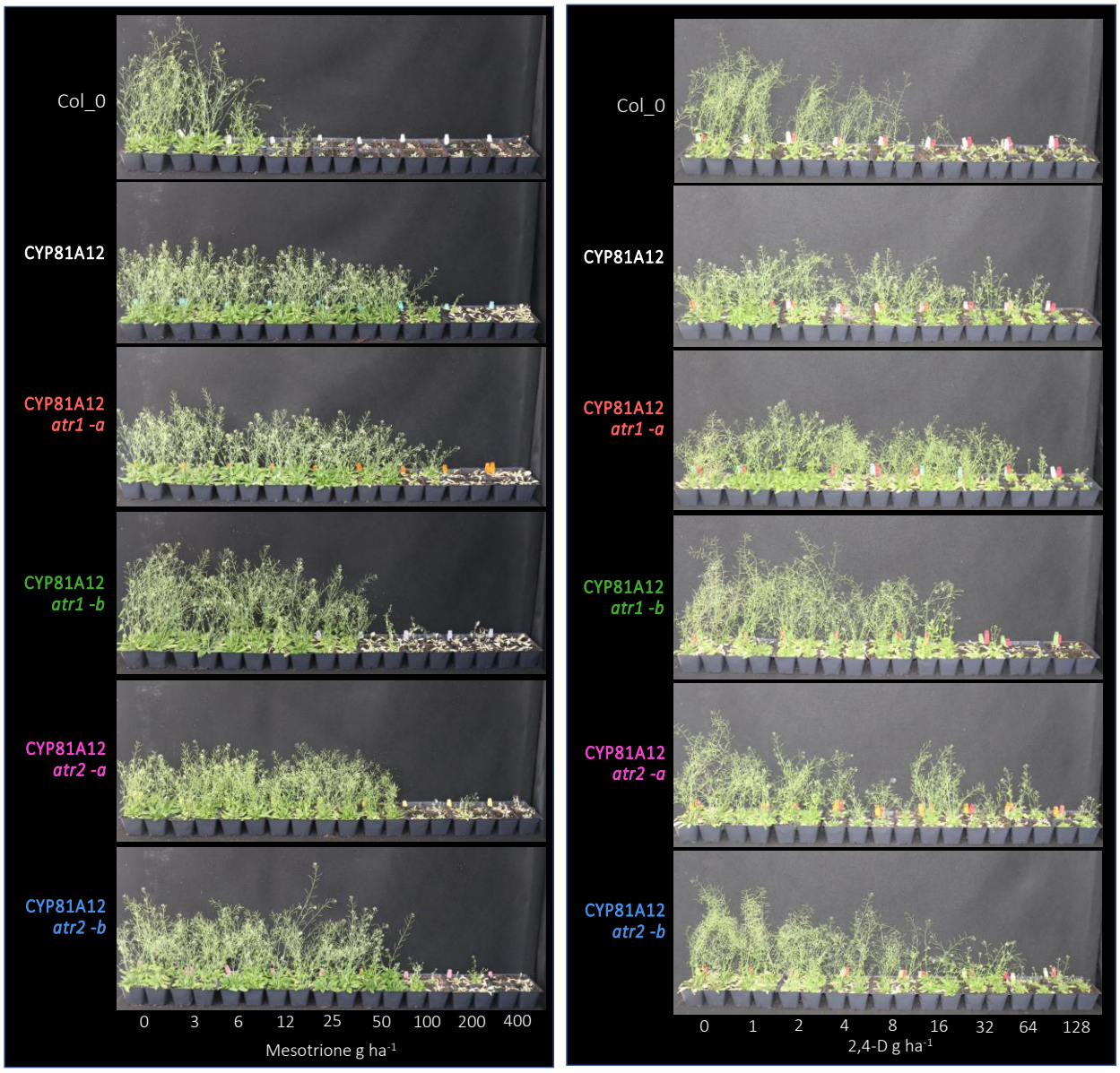
Supporting Information Table 4-2. Weed species and their information about NADPH-cytochrome P450 composition.

species	common name	ploidy	#chr	#CPR	# copies	CPR class	source
Alopecurus myosuroides	blackgrass	diploid	14	3	1	II	IWCG[1]
Amaranthus palmeri	Palmer amaranth	diploid	32	2	1	II	CoGe[2]
Amaranthus tuberculatus	waterhemp	diploid	32	2	1	I and II	CoGe
Bassia scoparia	kochia	diploid	9	2	1	I and II	IWCG
Bromus tectorum	downy brome	diploid	7	2	1	II	IWCG
Chenopodium formosanum	Djulis	hexaploid	27	2	3	I and II	IWCG
Conyza canadensis	horseweed	diploid	18	3	1	II	IWCG
Echinochloa colona	jungle rice	hexaploid	27	3	3	II	IWCG
Echinochloa crus-galli	barnyardgrass	hexaploid	27	3	2 and 3	II	IWCG
Echinochloa oryzicola	late watergrass	tetraploid	18	3	3	II	IWCG
Eleusine indica	goosegrass	diploid	9	3	1	II	IWCG
Ipomoea purpurea	common morning-glory	diploid	15	2	1 and 2	I and II	IWCG
Poa annua	annual bluegrass	tetraploid	28	3	2	II	IWCG

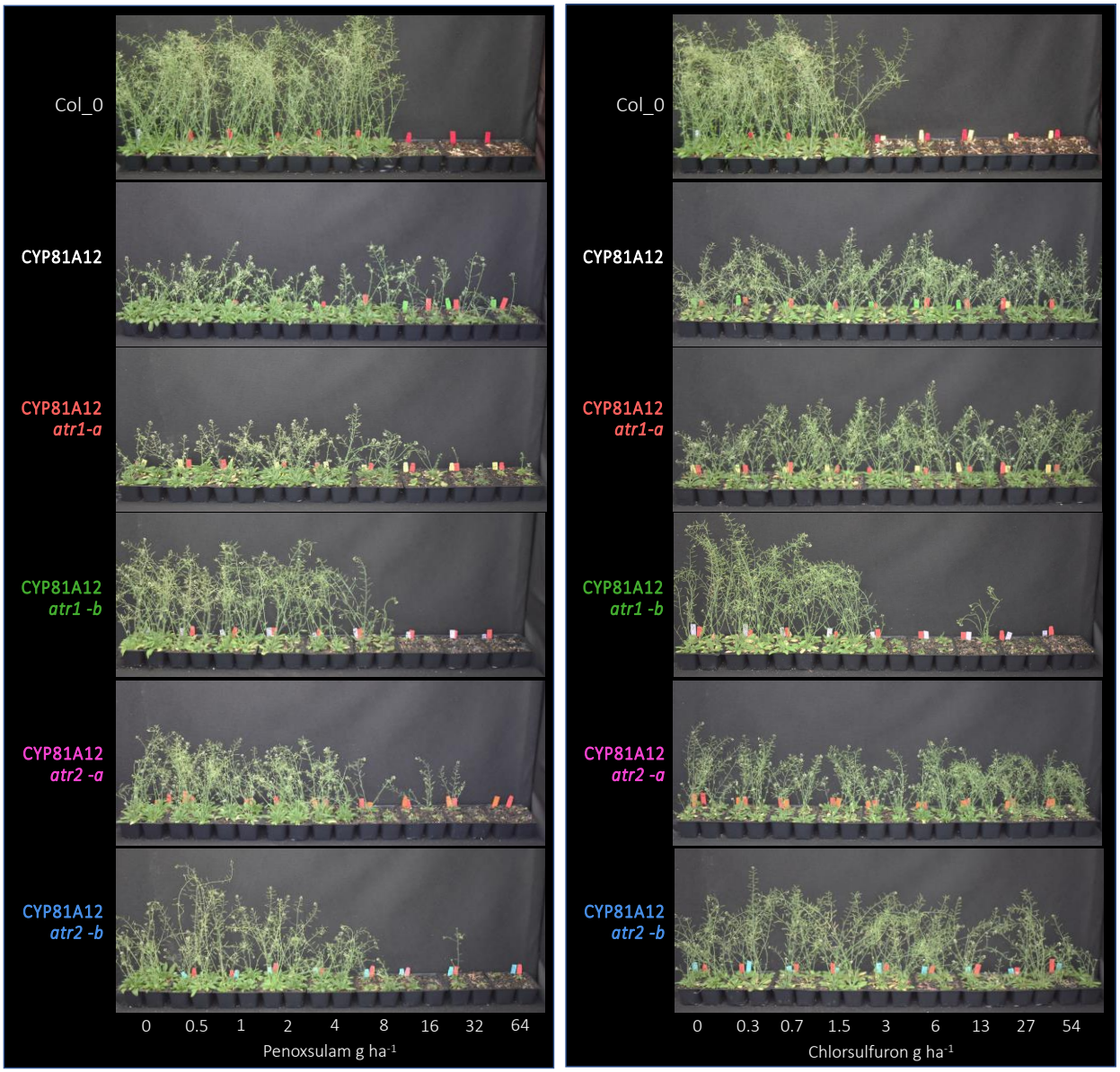
#chr – species chromosome number, #CPR – number of NADPH cytochrome P450 reductase, [1]IWCG – International Weed Genomic Consortium, [2]CoGe - platform for performing Comparative Genomics research.



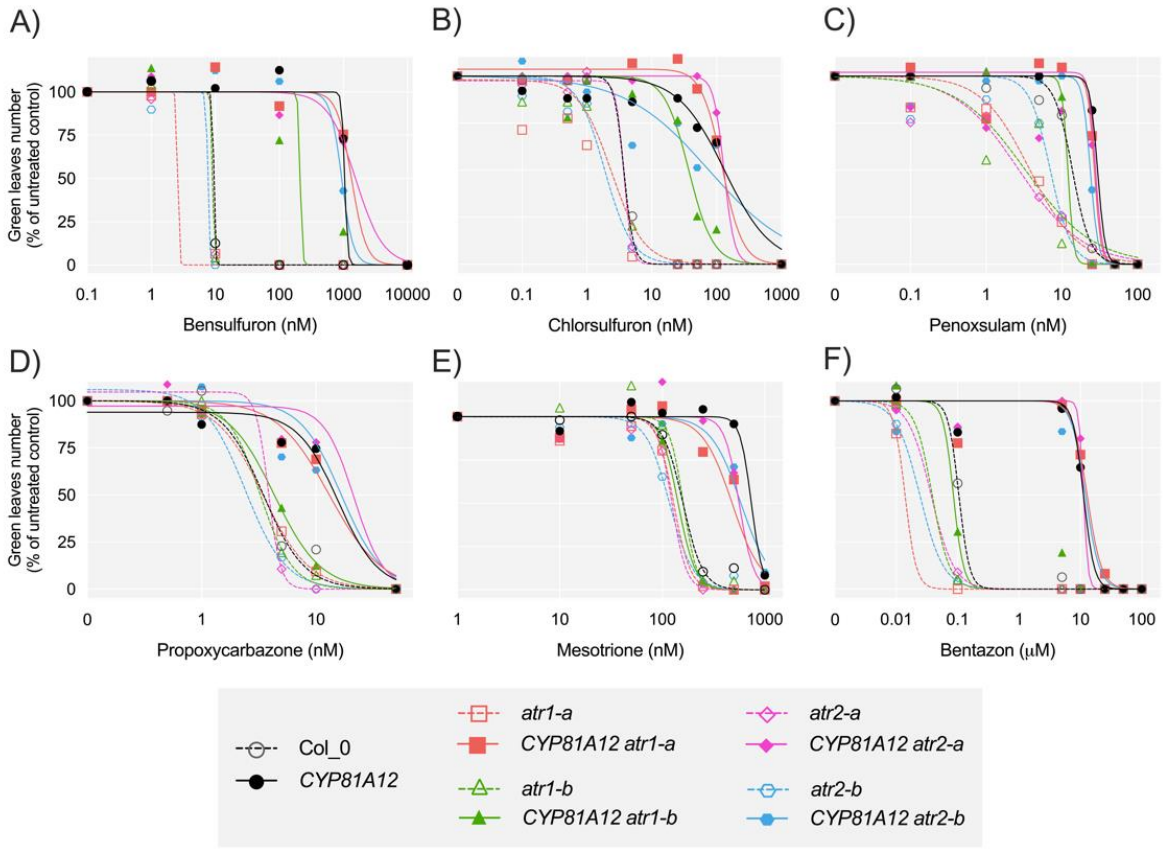
Supporting Information Figure 4-1**.** Phenotype analysis of transgenic mutant lines. A) Picture at 28 days after transplanting of four replicates for each line. B) Fresh mass of twelve plants of each line at 28 days after transplanting.



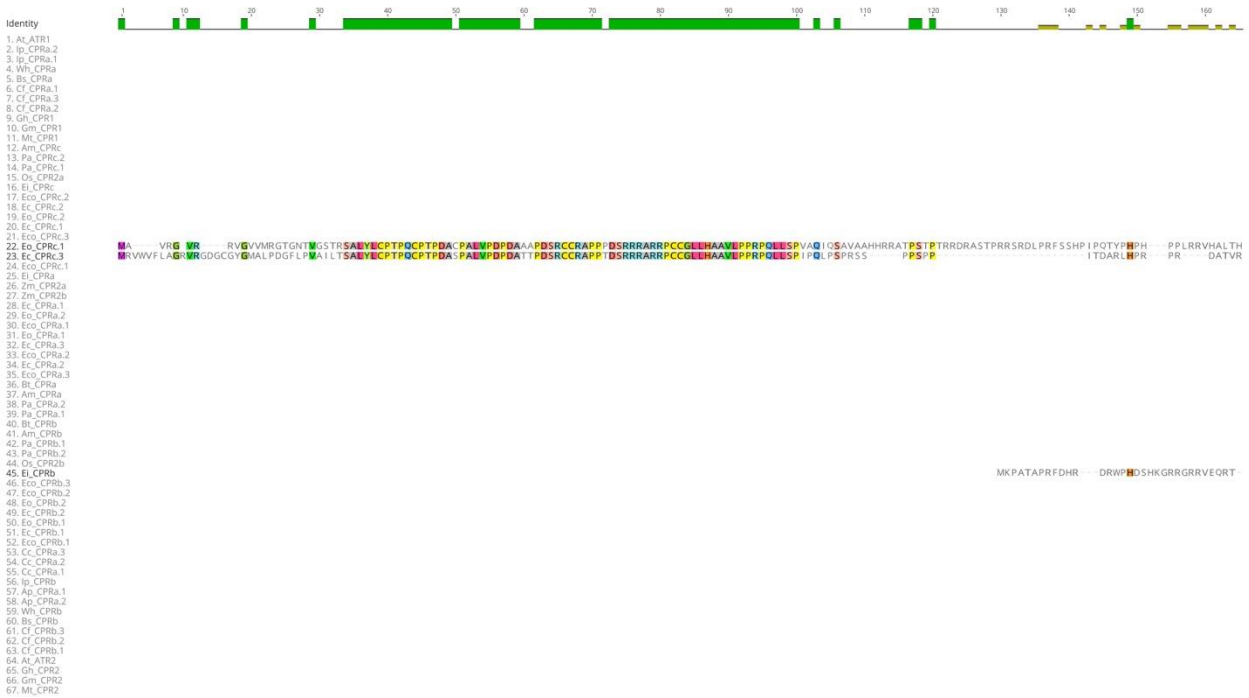
Supporting Information Figure 4-2. Pictures of different lines of Arabidopsis submitted to different doses of mesotrione and 2,4-D. Pictures of four replicates at 28 d after herbicide application. Col_0, wild-type; *CYP81A12*, transgenic line expression *CYP81A12*; *CYP81A12 atr1-a*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR1* 5'UTR; *CYP81A12 atr1-b*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR1* 4th intron; *CYP81A12 atr2-a*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR2* 3th intron; *CYP81A12 atr2-b*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR2* 12th exon.

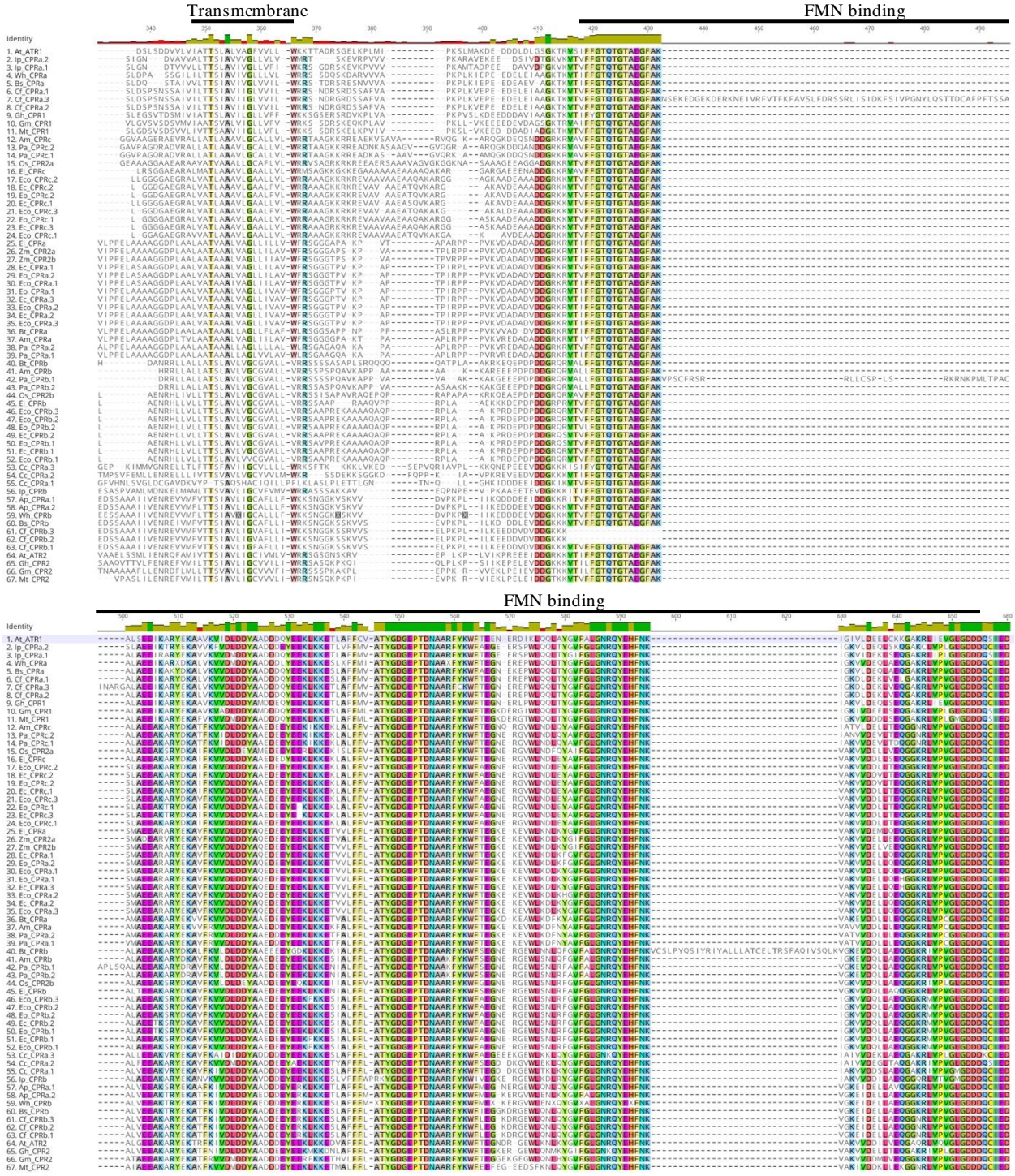


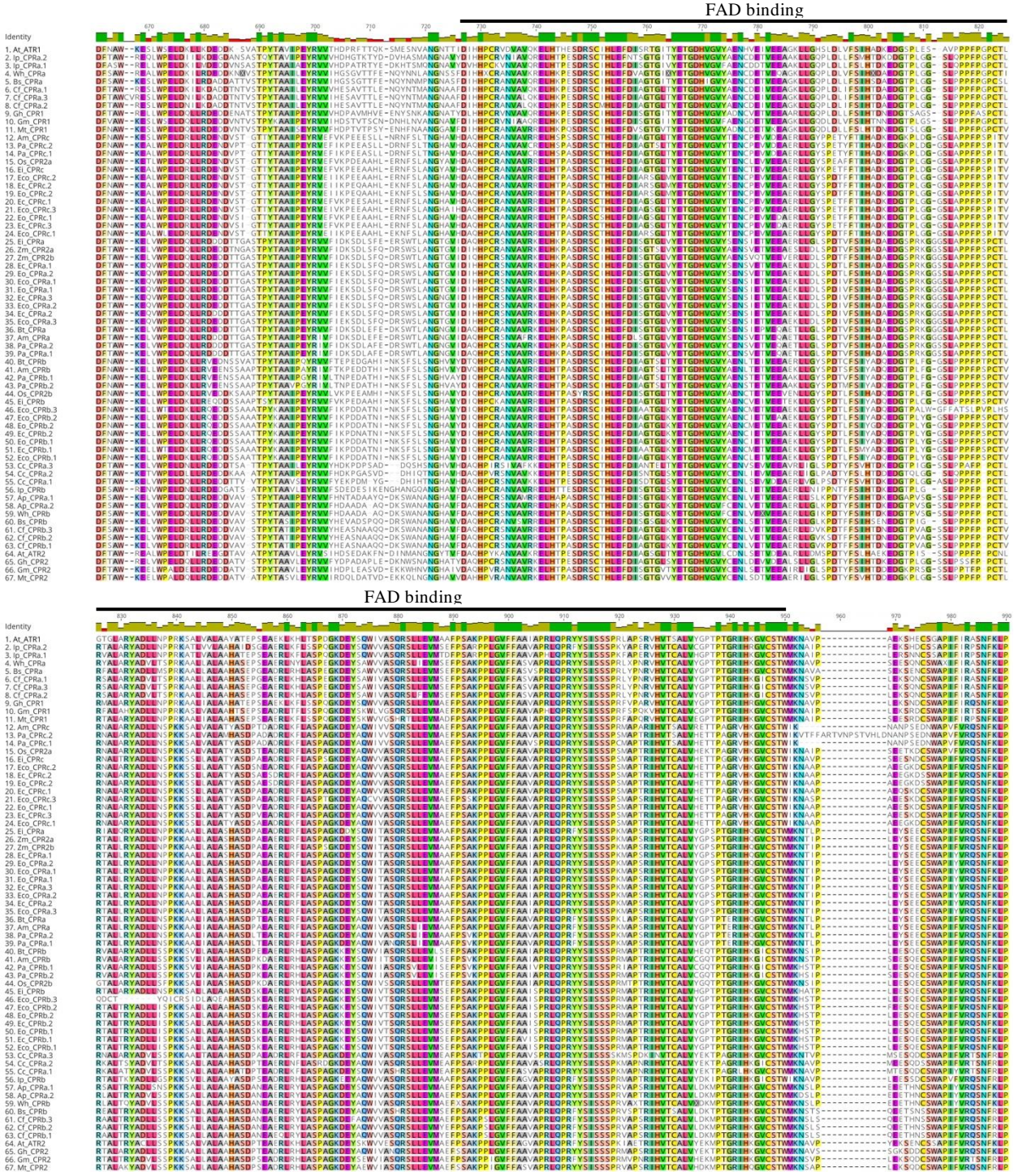
Supporting Information Figure 4-3. Pictures of different lines of Arabidopsis submitted to different doses of penoxsulam and chlorsulfuron. Pictures of four replicates at 28 d after herbicide application. Col_0, wild-type; *CYP81A12*, transgenic line expression *CYP81A12*; *CYP81A12 atr1-a*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR1* 5'UTR; *CYP81A12 atr1-b*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR1* 4th intron; *CYP81A12 atr2-a*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR2* 3th intron; *CYP81A12 atr2-b*, transgenic Arabidopsis expressing *CYP81A12* carrying T-DNA insertion on *ATR2* 12th exon.



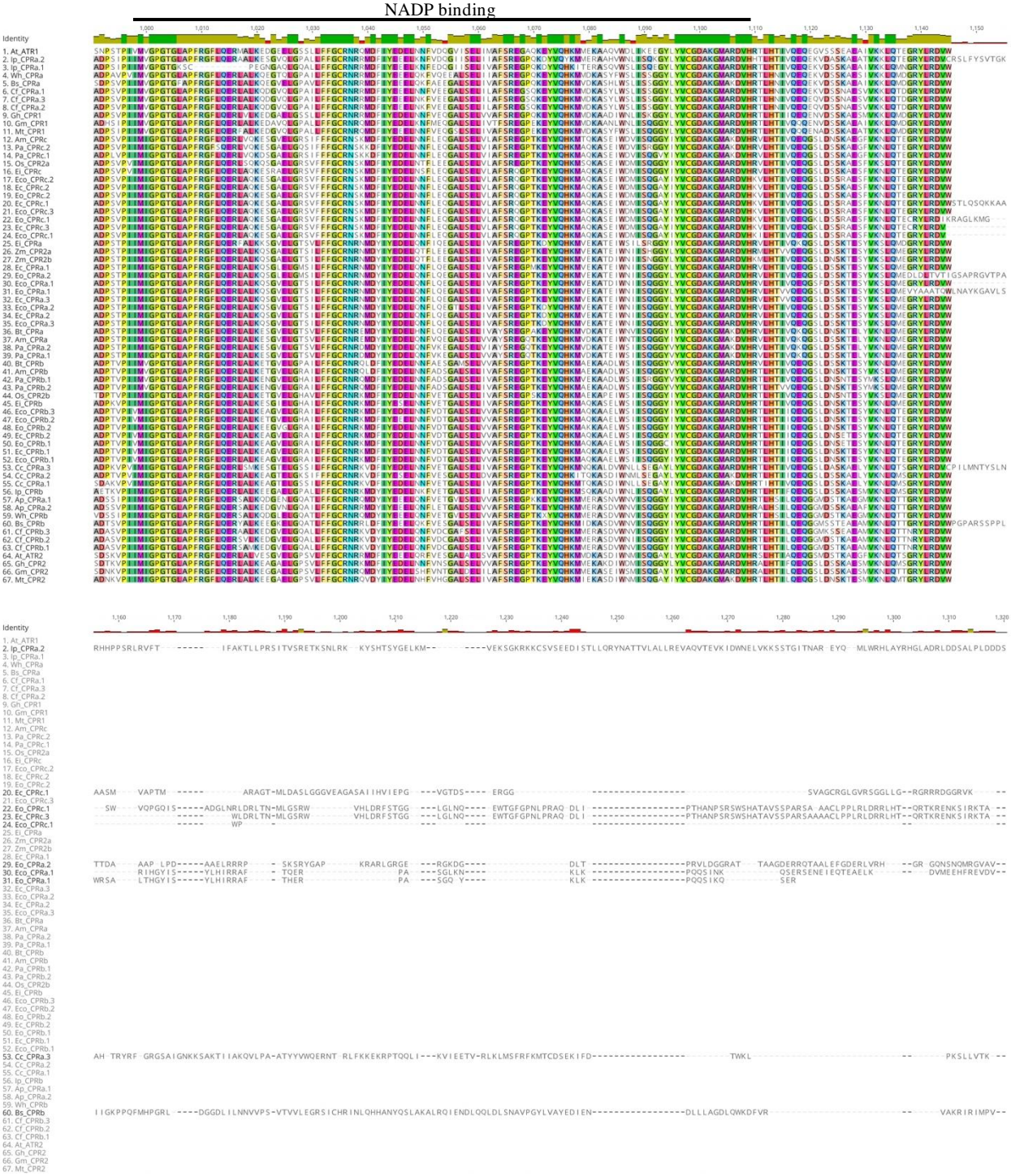
Supporting Information Figure 4-4. Dose-response of different lines of Arabidopsis growing in different herbicides concentration in ½ MS plates. Arabidopsis wild-type (Col_0), transgenic *CYP81A12*, mutant lines *atr1-a*, *atr1-b*, *atr2-a* and *atr2-b* and transgenic mutant lines expressing *CYP81A12, CYP81A12 atr1-a*, *CYP81A12 atr1-b*, *CYP81A12 atr2-a* and *CYP81A12 atr2-b*. A) dose-response curve to bensulfuron, B) chlorsulfuron, C) penoxsulam, D) propoxycarbazone, E) mesotrione and F) bentazon.

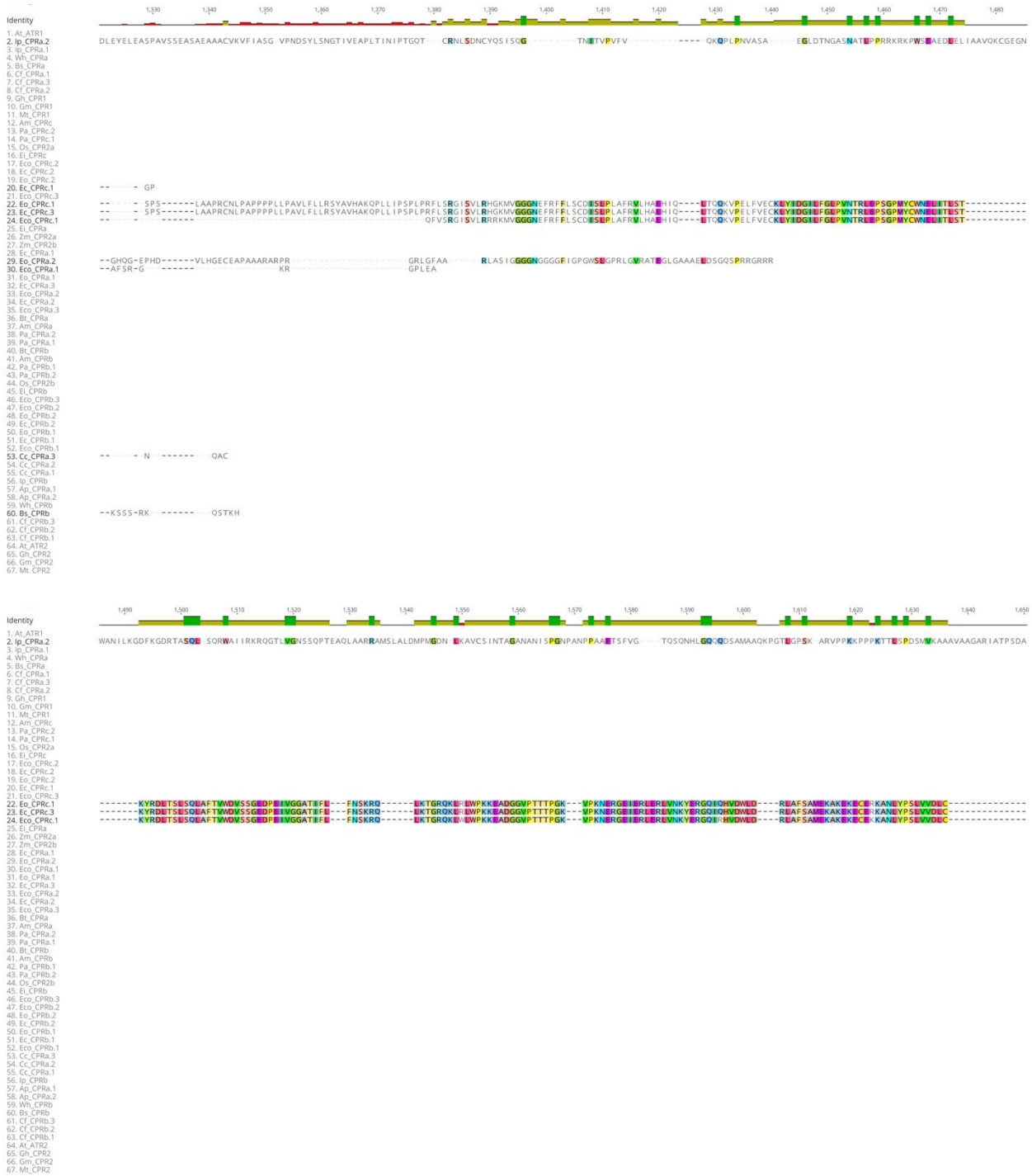

Identity
1. At_ATR1
2. Ip_CPRa.2
3. Ip_CPRa.1
4. Wh_CPRa
5. Bs_CPRa
6. Cf_CPRa.1
7. Cf_CPRa.3
8. Cf_CPRa.2
9. Gh_CPR1
10. Gm_CPR1
11. Mt_CPR1
12. Am_CPRc
13. Pa_CPRc.2
14. Pa_CPRc.1
15. Os_CPR2a
16. Ei_CPRc
17. Eco_CPRc.2
18. Ec_CPRc.2
19. Eo_CPRc.2
20. Ec_CPRc.1
21. Eco_CPRc.3
22. Eo_CPRc.1
23. Ec_CPRc.3
24. Eco_CPRc.1
25. Ei_CPRa
26. Zm_CPR2a
27. Zm_CPR2b
28. Ec_CPRa.1
29. Eo_CPRa.2
30. Eco_CPRa.1
31. Eo_CPRa.1
32. Ec_CPRa.3
33. Eco_CPRa.2
34. Ec_CPRa.2
35. Eco_CPRa.3
36. Bt_CPRa
37. Am_CPRa
38. Pa_CPRa.2
39. Pa_CPRa.1
40. Bt_CPRb
41. Am_CPRb
42. Pa_CPRb.1
43. Pa_CPRb.2
44. Os_CPR2b
45. Ei_CPRb
46. Eco_CPRb.3
47. Eco_CPRb.2
48. Eo_CPRb.2
49. Ec_CPRb.2
50. Eo_CPRb.1
51. Ec_CPRb.1
52. Eco_CPRb.1
53. Cc_CPRa.3
54. Cc_CPRa.2
55. Cc_CPRa.1
56. Ip_CPRb
57. Ap_CPRa.1
58. Ap_CPRa.2
59. Wh_CPRb
60. Bs_CPRb
61. Cf_CPRb.3
62. Cf_CPRb.2
63. Cf_CPRb.1
64. At_ATR2
65. Gh_CPR2
66. Gm_CPR2
67. Mt_CPR2

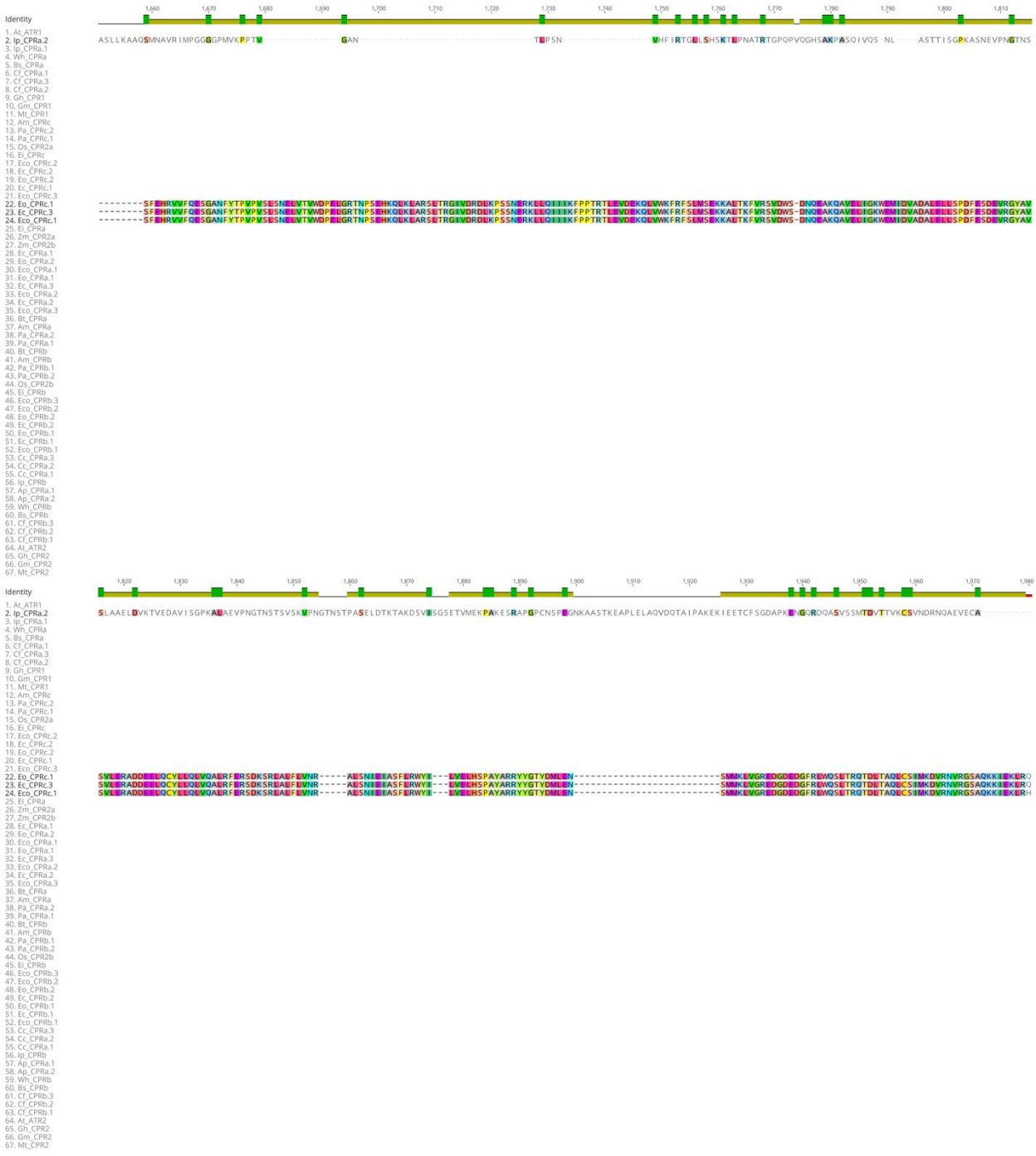

Transmembrane
FMN binding
FMN binding

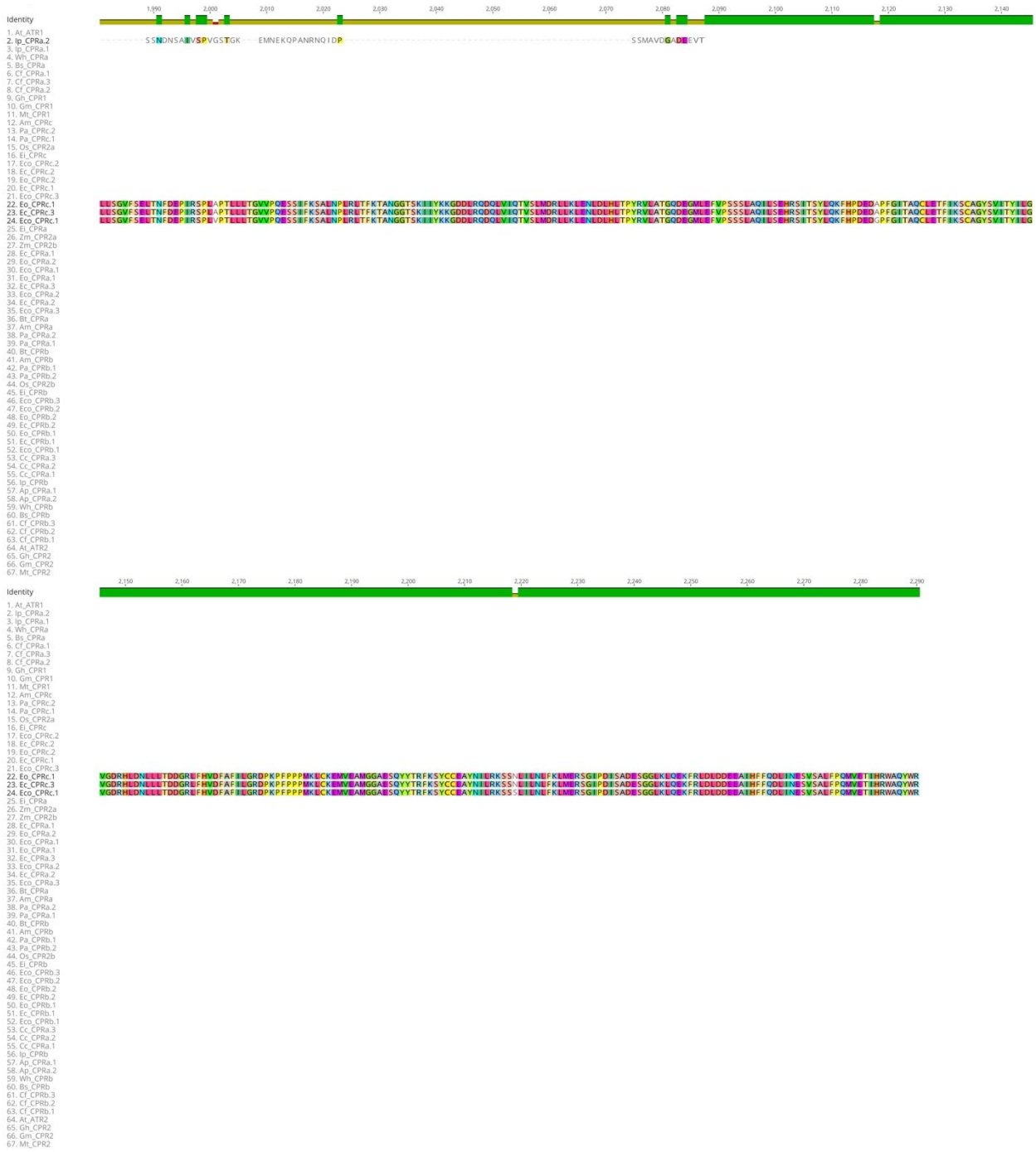

FAD binding
FAD binding

NADP binding




Identity
1. At_ATR1
2. Ip_CPRa.2
3. Ip_CPRa.1
4. Wh_CPRa
5. Bs_CPRa
6. Cf_CPRa.1
7. Cf_CPRa.3
8. Cf_CPRa.2
9. Gh_CPR1
10. Gm_CPR1
11. Mt_CPR1
12. Am_CPRc
13. Pa_CPRc.2
14. Pa_CPRc.1
15. Os_CPR2a
16. El_CPRc
17. Eco_CPRc.2
18. Ec_CPRc.2
19. Eo_CPRc.2
20. Ec_CPRc.1
21. Eco_CPRc.3
22. Eo_CPRc.1
23. Ec_CPRc.3
24. Eco_CPRc.1
25. El_CPRa
26. Zm_CPR2a
27. Zm_CPR2b
28. Ec_CPRa.1
29. Eo_CPRa.2
30. Eco_CPRa.1
31. Eo_CPRa.1
32. Ec_CPRa.3
33. Eco_CPRa.2
34. Ec_CPRa.2
35. Eco_CPRa.3
36. Bt_CPRa
37. Am_CPRa
38. Pa_CPRa.2
39. Pa_CPRa.1
40. Bt_CPRb
41. Am_CPRb
42. Pa_CPRb.1
43. Pa_CPRb.2
44. Os_CPR2b
45. El_CPRb
46. Eco_CPRb.3
47. Eco_CPRb.2
48. Eo_CPRb.2
49. Ec_CPRb.2
50. Eo_CPRb.1
51. Ec_CPRb.1
52. Eco_CPRb.1
53. Cc_CPRa.3
54. Cc_CPRa.2
55. Cc_CPRa.1
56. Ip_CPRb
57. Ap_CPRa.1
58. Ap_CPRa.2
59. Wh_CPRb
60. Bs_CPRb
61. Cf_CPRb.3
62. Cf_CPRb.2
63. Cf_CPRb.1
64. At_ATR2
65. Gh_CPR2
66. Gm_CPR2
67. Mt_CPR2


Identity
1. At_ATR1
2. Ip_CPRa.2
3. Ip_CPRa.1
4. Wh_CPRa
5. Bs_CPRa
6. Cf_CPRa.1
7. Cf_CPRa.3
8. Cf_CPRa.2
9. Gh_CPR1
10. Gm_CPR1
11. Mt_CPR1
12. Am_CPRc
13. Pa_CPRc.2
14. Pa_CPRc.1
15. Os_CPR2a
16. Ei_CPRc
17. Eco_CPRc.2
18. Ec_CPRc.2
19. Eo_CPRc.2
20. Ec_CPRc.1
21. Eco_CPRc.3
22. Eo_CPRc.1
23. Ec_CPRc.3
24. Eco_CPRc.1
25. Ei_CPRa
26. Zm_CPR2a
27. Zm_CPR2b
28. Ec_CPRa.1
29. Eo_CPRa.2
30. Eco_CPRa.1
31. Eo_CPRa.1
32. Ec_CPRa.3
33. Eco_CPRa.2
34. Ec_CPRa.2
35. Eco_CPRa.3
36. Bt_CPRa
37. Am_CPRa
38. Pa_CPRa.2
39. Pa_CPRa.1
40. Bt_CPRb
41. Am_CPRb
42. Pa_CPRb.1
43. Pa_CPRb.2
44. Os_CPR2b
45. Ei_CPRb
46. Eco_CPRb.3
47. Eco_CPRb.2
48. Eo_CPRb.2
49. Ec_CPRb.2
50. Eo_CPRb.1
51. Ec_CPRb.1
52. Eco_CPRb.1
53. Cc_CPRa.3
54. Cc_CPRa.2
55. Cc_CPRa.1
56. Ip_CPRb
57. Ap_CPRa.1
58. Ap_CPRa.2
59. Wh_CPRb
60. Bs_CPRb
61. Cf_CPRb.3
62. Cf_CPRb.2
63. Cf_CPRb.1
64. At_ATR2
65. Gh_CPR2
66. Gm_CPR2
67. Mt_CPR2



Supporting Information Figure 4-5. Multiple protein alignment of 67 CPR protein sequences from 13 different weed species and others six plant species with colors indicating conserved residue between sequences. Transmembrane (347-366), FMN binding (418-665), FAD binding (727-950) and NADP binding (999-1109) are indicated. Different CPR within species were named as “a”, “b” or “c” and the copies were numbered as “.1”, “.2” or “.3” for one, two or three copies, respectively. Am - *Alopecurus myosuroides* (blackgrass), At - *Arabidopsis thaliana* (Arabidopsis), Ap – *Amaranthus palmeri* (Palmer amaranth), Wh – *Amaranthus tuberculatus* (waterhemp), Bs - *Bassia scoparia* (Kochia), Bt - *Bromus tectorum* (downy brome), Cc - *Conyza canadensis* (horseweed), Cf - *Chenopodium formosanum* (rainbow rice), Ec - *Echinochloa crus-*

galli (barnyardgrass), *Eo* - *E. oryzicola* (late watergrass), Eco - *E. colona* (jungle rice), Ei - *Eleusine indica* (goosegrass), Gh - *Gossypium hirsutum* (cotton), Gm - *Glycine max* (soybean), Ip - *Ipomoea purpurea* (common morning glory), Mt - *Medicago truncatula* (Barrel medic), Os - *Oryza sativa* (rice), Pa - *Poa annua* (annual bluegrass).