Investigating glyphosate tolerance mechanisms in the liverwort *Marchantia polymorpha*



Samuel George Caygill

Hertford College
Department of Biology
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy
Michaelmas 2023

Investigating glyphosate tolerance mechanisms in the liverwort *Marchantia polymorpha*

Samuel Caygill

Hertford College, Department of Biology, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Michaelmas 2023

Abstract

The evolution of herbicide resistance in weeds is a threat to global crop yields. The aim of this thesis is to increase our understanding of the evolution and mechanisms of herbicide resistance. To achieve this, I first determine the phylogenetic relationships of both the ATP-binding cassette (ABC) transporter and uridine diphosphate glycosyltransferase (UGT) protein superfamilies within the Archaeplastida lineage. Proteins from these families confer non-target site resistance (NTSR) in populations of weeds. I also determine that the liverwort *Marchantia polymorpha* is naturally tolerant to the non-selective herbicide glyphosate. Using enzyme activity assays, I demonstrate that glyphosate tolerance in *M. polymorpha* is not due to a resistant version of the glyphosate target EPSPS. Instead, I discover a novel form of glyphosate tolerance conferred by the peptidoglycan biosynthesis enzyme MurA, which is structurally similar to EPSPS. By expressing MurA in the angiosperm *Arabidopsis thaliana* and generating *murA* loss-of-function mutants in *M. polymorpha*, I determine that MurA confers glyphosate tolerance independently from its function in peptidoglycan biosynthesis. I show that MpMurA catalyses the same reaction as EPSPS, transferring the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to shikimate-3-phosphate (S3P) forming 5-enolpyruvylshikimate-3-phosphate (EPSP) product. However, like EPSPS, it is also inhibited by glyphosate. I therefore propose that MurA confers glyphosate tolerance either through a mechanism analogous to target-site overexpression or by binding glyphosate, making it less available to inhibit EPSPS. Through a metabolomic analysis, I also show that the metabolic response to glyphosate treatment in *M. polymorpha* differs from that in *A. thaliana*. This may be a result of, or contribute to, glyphosate tolerance in *M. polymorpha*. These results therefore demonstrate a novel form of glyphosate tolerance and determine the physiological impacts of glyphosate treatment that may contribute to its mode of action.

Contents

Acknowledgements 7

Abbreviations 8

Chapter 1: Introduction 11

1.1. Summary 12

1.2. Herbicides are an essential tool in intensive modern agriculture 13

1.3. The evolution of herbicide resistance in weeds 14

1.4. Mechanisms of herbicide resistance 17
1.4.1. Target-site resistance (TSR) 18
1.4.2. Non-target site resistance (NTSR) 18

1.5. NTSR: current knowledge 19
1.5.1. Cytochrome P450s (P450s) 20
1.5.2. Glutathione S-transferases (GSTs) 21
1.5.3. Uridine diphosphate (UDP) glycosyltransferases (UGTs) 22
1.5.4. ATP binding cassette (ABC) transporters 23
1.5.5. Reduced translocation and rapid cell death 24
1.5.6. NTSR summary 24

1.6. Glyphosate and glyphosate resistance 25

1.7. Using *Marchantia polymorpha* to study glyphosate resistance 30

Chapter 2: ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation 35

2.1 Abstract 37

2.2 Introduction 38

2.3 Materials and Methods 42
2.3.1 Data Resources 42
2.3.2 Sequence collection 43
2.3.3 Sequence alignment 43
2.3.4 Phylogenetic analysis 43

2.4 Results 44
2.4.1 960 ABC transporter and 552 UGT sequences were identified in the genomes of the 10 species of Archaeplastida 44
2.4.2 Plant ABC transporters are ancient, and all 12 groups existed in the last common ancestor of the Archaeplastida 46
2.4.4 Six ABC-transporter sub-families have been associated with non-target site resistance to herbicides 56
2.4.5 UGT groups associated with herbicide resistance evolved after the divergence of Angiosperms from the Archaeplastida lineage 60

2.5 Discussion 64

2.6 Supplementary Data 68

Chapter 3: *Marchantia polymorpha* is glyphosate-tolerant but lacks a glyphosate-tolerant form of the EPSPS enzyme 71

3.1. Abstract 72

3.2. Introduction 73

3.3. Material and Methods 75
3.3.1. Phylogenetic analysis of EPSPS genes 75
3.3.2. Homology modelling 75
3.3.3. Molecular docking 76
3.3.4. Plant lines and growth conditions 76
3.3.5 DNA constructs 76
3.3.6. *Agrobacterium tumefaciens*-mediated transformation of *A. thaliana* plants 79
3.3.7. Herbicide and stock solution preparation 79
3.3.8. Dose-response assays 79
3.3.9. Site-directed mutagenesis 80
3.3.10. Crude-EPSPS isolation 80
3.3.11. EPSPS expression and purification 81
3.3.12. EPSPS enzyme assays 81
3.3.13. Primers 82
3.3.14. Buffers and Media 83

3.4. Results 84
3.4.1. *M. polymorpha* plants are more tolerant to application of glyphosate than *A. thaliana* plants 84
3.4.2. The EPSPS protein sequence and structure is highly conserved among land plants, including *M. polymorpha* 86
3.4.3. Glyphosate is modelled to bind to the EPSPS of *M. polymorpha* in the same confirmation as to a glyphosate-resistant EPSPS enzyme 89
3.4.4. *M. polymorpha* crude protein extract containing EPSPS is not more tolerant to glyphosate than crude protein extracts containing EPSPS from other land plants 92
3.4.5. Purified *M. polymorpha* EPSPS is sensitive to glyphosate 94
3.4.6. The kinetic properties of plant EPSPSs are similar to each other but different from bacterial EPSPS enzymes 96
3.4.7. Overexpression of the *A. thaliana EPSPS* gene leads to higher glyphosate resistance than overexpression of the *M. polymorpha EPSPS* gene 99
3.4.8. *In silico* analysis of *A. thaliana* and *M. polymorpha* reveals a residue in *M. polymorpha* that may lead to differences in kinetic parameters 102
3.4.9. A serine to valine mutation at position 500 in the *M. polymorpha* EPSPS does not increase glyphosate sensitivity *in planta* 104
3.4.10. A valine to serine mutation at position 493 in the *A. thaliana* EPSPS does not increase glyphosate resistance *in planta* 106

3.5. Discussion 108

3.6. Supplementary data 113

Chapter 4: The peptidoglycan biosynthesis enzyme UDP-N-acetylglucosamine-3-enopyruvyl transferase (MurA) confers glyphosate tolerance in *Marchantia polymorpha* 121

4.1 Abstract 122

4.2. Introduction 123

4.3. Materials and Methods 126
4.3.1. Protein Structure Prediction using AlphaFold 126
4.3.2. Phylogenetic analysis of *MurA* genes 126
4.3.3. *M. polymorpha* spore sterilisation 126
4.3.4. Plant lines and growth conditions 126
4.3.5. Construct making 126
4.3.6. Guide RNA design 128
4.3.7. Cloning of CRISPR-Cas9 vectors 128
4.3.8. *Agrobacterium tumefaciens*-mediated transformation of *M. polymorpha* spores 129
4.3.9. Agrobacterium tumefaciens-mediated transformation of *A. thaliana* plants 130
4.3.10. Genotyping putative mutants generated by CRISPR/Cas9 mutagenesis 130
4.3.11. Phenotyping putative mutants generated by CRISPR/Cas9 mutagenesis and plants treated with peptidoglycan biosynthesis inhibiting antibiotics 130

4.3.12. Herbicide and stock solution preparation 130
4.3.13. Dose-response assays 130
4.3.14. MurA and EPSPS expression and purification 131
4.3.15. MurA and EPSPS enzyme assays 131
4.3.16. Micro-scale thermophoresis binding assays 131
4.3.17. Primers 133
4.3.18. Media and buffers 134

4.4. Results 135
4.4.1. The enzyme MurA is encoded by *Marchantia polymorpha* and is structurally similar to EPSPS 135
4.4.2. EPSPS is conserved throughout the Archaeplastida, but MurA is absent from Angiosperms 137
4.4.3. Land plants that encode a MurA enzyme are tolerant to glyphosate 140
4.4.4. Overexpression of Mp*MurA* confers increased resistance to glyphosate 142
4.4.5. Seven predicted Mp*murA* loss-of-function mutants were generated using CRISPR-Cas9 mutagenesis 144
4.4.6. Predicted Mp*murA* loss-of-function mutants are more sensitive to glyphosate than wild-type plants 150
4.4.7. Five predicted Mp*murF* loss-of-function mutants were generated using CRISPR-Cas9 mutagenesis 152
4.4.8. Predicted Mp*murF* loss-of-function mutants are not more sensitive to glyphosate than wild-type plants 158
4.4.9. Treatment of *M. polymorpha* with some peptidoglycan biosynthesis inhibiting antibiotics leads to a macrochloroplast phenotype 160
4.4.10. Treatment of *M. polymorpha* with peptidoglycan biosynthesis inhibiting antibiotics does not alter sensitivity to glyphosate 165
4.4.11. Expression of the Mp*MurA* gene confers glyphosate resistance in *A. thaliana* 167
4.4.12. The overexpression of partial forms of MpMurA in *M. polymorpha* does not increase glyphosate resistance 169
4.4.13. Micro-scale thermophoresis/Temperature -related intensity change analysis cannot detect glyphosate binding to MpMurA 172
4.4.14. MpMurA catalyses the reaction between shikimate-3-phosphate and phosphoenolpyruvate to produce 5-enolypyruvylshikimate-3-phosphate 174

4.4. Discussion 177

4.6. Supplementary data 185

Chapter 5: A widespread but distinct metabolomic response occurs following glyphosate-treatment in *Marchantia polymorpha* 190

5.1. Abstract 191

5.2. Introduction 192

5.3. Materials and methods 195
5.3.1 Plant materials and growth conditions 195
5.3.2 Growth assays with media supplemented with aromatic amino acids 195
5.3.3 Sensitivity of *M. polymorpha* and *A. thaliana* to shikimate 195
5.3.4 Measuring shikimate accumulation in plant tissue 195
5.3.5 Sample extraction for metabolomic analysis 196
5.3.6 Targeted metabolomic analysis using liquid chromatograph-mass spectrometry 196
5.3.7 Non-targeted metabolomic analysis using liquid chromatography-mass spectrometry 199
5.3.8 RNA-Seq experimental design and RNA extraction and sequencing 199
5.3.9 Read pre-processing and mapping to the reference transcriptome 200
5.3.10 Differential expression analysis 200

5.4. Results 201
5.4.1 Glyphosate is taken up in larger quantities in *M. polymorpha* than *A. thaliana* but diminishes at a faster rate 201

5.4.2 Shikimate accumulates more upon glyphosate treatment in the glyphosate sensitive *A. thaliana* than the glyphosate-tolerant *M. polymorpha* 203
5.4.3 Exogenous application of shikimate reduces growth of *A. thaliana* but not *M. polymorpha* 206
5.4.4 Glyphosate lethality in *A. thaliana* can be rescued with the co-treatment of aromatic amino acids 208
5.4.5 Pools of amino acids are affected differently in *A. thaliana* compared to *M. polymorpha* 210
5.4.6 Quantities of intermediate metabolites of the shikimate pathway are differently affected by glyphosate treatment in *M. polymorpha* and *A. thaliana* 213
5.4.7 Glyphosate treatment causes widespread, but different, metabolic changes in *M. polymorpha* and *A. thaliana* 216
5.4.8 Many of the metabolic pathways affected by glyphosate treatment are common between *M. polymorpha* and *A. thaliana* 219
5.4.9 There are large and opposite impacts on the TCA cycle in *A. thaliana* and *M. polymorpha* after 7 days of glyphosate treatment 221
5.4.10 Many intermediates of glycolysis decrease in *M. polymorpha* upon glyphosate treatment but increase in *A. thaliana* 224
5.4.11 There are very few transcriptome changes in *M. polymorpha* following glyphosate treatment 227

5.5. Discussion 229

5.6. Supplemetary data 238

Chapter 6: General Discussion 239

Appendix 246

References 270

Acknowledgements

First, I would like to say a big thanks to my supervisor Prof. Liam Dolan for your advice, support, and mentorship throughout my DPhil. I am very grateful of the freedom you allowed me in my project, letting me pursue it wherever it took me. To Prof. Gail Preston, Prof. Philip Poole, Prof. Renier Van der Hoorn, and Prof. Nick Harberd, thank you also for your advice and guidance. Many thanks to the BBSRC and the Gregor Mendel Institute for funding me.

I would like to extend a massive thank you to everyone in the Dolan lab, past and present, for making such a positive working environment and for being so supportive and helpful. Special thank you to Dr. Alex Casey and Dr. Chloe Casey for helping with everything herbicide related and beyond, and for making the lockdown times in Vienna not so bad. To Dr. Eva-Sophie Wallner for your help and patience teaching me various methods and getting me started with *Arabidopsis*. To Dr. Hugh Mulvey, Dr. Victoria Spencer, Dr. Zohar Meir, Sarah Attrill, and Johannes Rötzer, thank you so much for you input of suggestions and ideas during lab meetings and lunches, which were a huge help to my project. Thank you very much to Dr. Magda Mosiolek and Kathi Jandrasits for supporting us all in the lab and without whom my project would not have been feasible. Thank you to all of the Vienna Biocenter Core Facilities for your enthusiastic help and work to contribute to my project.

Finally, I would like to thank all my friends and family. To Sasha and Georgia, for making my time in Vienna so wonderful and always being there to talk and to Sasha for giving me respite from work on our weekly lunch. A big shoutout to all the boys, thanks for your visits to Vienna and all the wholesome times we've had and will have. Thanks to all the other wonderful friends I've made in Vienna who make life here so great and have got me to the end of my DPhil. Of course, to Frida, I'm very grateful to have so serendipitously met you and for always giving me things to look forward to, thank you. And finally, a big thank you to Mum and Dad for everything you do for me, I truly am grateful!

Abbreviations

2,4-D	2,4-dichlorophenoxyacetic acid
AAA	Aromatic Amino Acid
ABC transporter	ATP-binding Cassette Transporter
Accase	Acetyl-CoA Carboxylase
aLRT	Approximate Likelihood-Ratio Test
ALS	Acetolactate Synthase
AMPA	aminomethylphosphonic acid
ANOVA	Anaysis of Variance
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base Pair
CaMV	Cauliflower Mosaic Virus
Col-0	Columbia-0
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DCA	Dichloroaniline
DHAP	Dihydroxycetone Phosphate
DNA	Deoxyribonucleic Acid
EPSP	5-enolpyruvylshikimate-3-phosphate
EPSPS	5-enolpyruvylshikimate-3-phosphate Synthase
FW	Fresh Weight
GOF	Gain of Function
GR	Glyphosate Resistant
GSH	Glutathione
GST	Glutathione S-transferase
GT	Glycosyltransferase
GWAS	Genome Wide Association Study
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High-Performance Liquid Chromatography
HRAC	Herbicide Resistance Action Committee
IAA	Indole-3-acetic acid

IC_{50}	Half Maximal Inhibitory Concentration
IMAC	Immobilised Metal Affinity Chromatography
IR	Infra-Red
K_{cat}	Catalytic Constant
K_{cat}/K_m	Catalytic Efficiency
K_D	Dissociation Constant
K_m	Michaelis Constant
LC-MS/MS	Liquid Chromatography tandem Mass Spectrometry
LOF	Loss of Function
ML	Maximum Likelihood
MOA	Mode of Action
Mp*murA*	*Marchantia polymorpha murA* mutant
Mp*murF*	*Marchantia polymorpha murF* mutant
MSMA	Monosodium Methanearsonate
MST	Micro-scale Thermophoresis
MWCO	Molecular Weight Cut-OFF
NADCC	Sodium Dichloroisocyanurate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NBD	Nucleotide Binding Domain
NTSR	Non-Target Site Resistance
P450	Cytochrome P450
PAM	Protospacer Adjacent Motif
PCG	Protein Coding Genes
PCR	Polymerase Chain Reaction
PEP	Phosphoenolpyruvate
PSII	Photosystem II
PSPG	Plant Secondary Product Glycosyltransferase
PT	Post-treatment
PVP	Polyvinylpyrrolidinone
RBCS	Rubisco Small Subunit
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid

RNA-seq	RNA sequencing
ROS	Reactive Oxygen Species
RP	Reversed Phase Chromatography
S3P	Shikimate-3-phosphate
SE	Standard Error
sgRNA	Single Guide RNA
Tak-1	Takaragaike-1
Tak-2	Takaragaike-2
TCA	Tricarboxylic Acid Cycle
TCP	Trichlorophenol
TIPS	T102I/P106S mutation
TMD	Transmembrane Domain
TRIC	Temperature Related Intensity Change
TSR	Target-Site Resistance
UDP	Uridine Diphosphate
UDP-GlcNAc	UDP N-acetlyglucosamine
UGT	UDP Glycosyltransferase
V_{max}	Maximum Velocity
WT	Wild-Type
YFP	Yellow Fluorescent Protein

Chapter 1: Introduction

1.1. Summary

In this introduction and literature review section, I outline the importance of herbicides in agriculture. I then introduce the herbicide resistance problem which has come about due to excessive use of a few herbicide modes of action, and which threatens global food production. I describe target site resistance (TSR) and non-target site resistance (NTSR) and explain the differences between these two main classes of herbicide resistance mechanisms. NTSR is more complex than TSR and I describe in detail the four major gene families which have most frequently been associated with herbicide resistance evolution in weed populations in the field. I also describe other mechanisms of NTSR which do not involve members of these four protein families.

I then introduce glyphosate, with the largest market share of all herbicides, and with many cases of herbicide resistance against it. I outline glyphosate's mode of action before reviewing the literature on known glyphosate resistance mechanisms. I highlight the severity of glyphosate resistance before outlining the need for efficient systems in which to study glyphosate resistance. I then introduce *M. polymorpha* as a system in which to study glyphosate resistance mechanisms, highlighting its benefits for herbicide and particularly glyphosate resistance research in the laboratory. I then discuss the influence of peptidoglycan, which is predicted to surround the chloroplast wall in *M. polymorpha*, on my research. I finally conclude my introduction with an outline of my DPhil research findings which investigated glyphosate tolerance mechanisms in *M. polymorpha*.

1.2. Herbicides are an essential tool in intensive modern agriculture

Plants are defined as weeds if they cause economic losses or ecological damage, create health problems for humans or animals or are considered undesirable where they are growing (WSSA, 1998). In agricultural systems, weeds compete with cultivated crops for essential resources such as light, water and nutrients, leading to potential reductions in crop yield and quality. This competition makes weeds the single greatest biotic threat to global crop yields, surpassing the collective impact of insect pests and plant pathogens (Oerke, 2006). Herbicides can mitigate the problems caused by weeds by targeting and reducing weed numbers and reducing their negative impact on crop yields. By minimising crop losses to weeds, herbicides can help to fulfil food production needs on the minimum footprint of land possible (Phalan et al., 2011). Furthermore, herbicides can support a number of sustainable farming practices such as no-till farming, which can reduce soil erosion, environmental degradation, and the amount of carbon dioxide in the atmosphere compared to farming with tilling (Friedrich, 2005; Morrison et al., 1990). Sustainable farming practices encompass those which allow us to meet our current food production needs without compromising the ability of future generations to meet their own needs (Lichtfouse et al., 2009).

Herbicides are chemicals which kill target plants by disrupting at least one vital physiological process. This usually involves the herbicide small molecule binding to a target within the plant and inhibiting its function. Before the widespread adoption of herbicide use in the developed world in the mid-20th century, weeds were controlled with a combination of manual weeding which is labour intensive, mechanical tillage which is energy intensive, and crop rotations (Harker and O'Donovan, 2013). Although some of these methods are still in use, the introduction of chemical herbicides to farmer's weed control arsenal provided an efficient and cost-effective solution (Heap, 2014). For these reasons, the use of herbicides is now increasing in the developing world (Gianessi, 2013; Hossain, 2015).

Herbicides are classified by their mode of action (MOA) and there are currently 26 herbicide MOAs on the market (Beffa et al., 2019; Heap, 2023). The MOA of a herbicide is defined by the cellular process which it inhibits and the herbicide MOAs currently available are

summarised by the Herbicide Resistance Action Committee (HRAC, 2022). Herbicides inhibit cellular processes by binding to a specific protein and inhibiting its function. The particular protein the herbicide binds to is referred to as the 'site of action' of the herbicide. There are currently 31 known sites of action of herbicides (Heap, 2023) Different herbicides are effective against different types of weeds and so this contributes to a demand for a variety of herbicides (Abbas et al., 2009; Hoss et al., 2003).

The types of herbicides and their application methods vary across different regions. This is for a variety of different reasons including availability, crops grown, regulatory policies, and types of weeds threatening crop yields due to geographical locality. For example, in the southern USA, Palmer Amaranth (*Amaranthus palmeri*) is a prevalent weed in cotton, soybean and maize, and is most often controlled with glyphosate (Ward et al., 2013). Blackgrass (*Alopecurus myosuroides*) is the most problematic weed in wheat fields of western Europe and largely controlled with flufenacet and pendimethalin (Hull and Moss, 2012; Moss et al., 2007). In Australia the weedy annual ryegrass (*Lolium multiflorum*) is controlled with a 'double knockdown' strategy of glyphosate followed by paraquat (Borger and Hashem, 2007). Therefore, the herbicide strategies required in different regions and for different farming types has driven the development and use of a variety of herbicide MOAs.

1.3. The evolution of herbicide resistance in weeds

Intensive use of herbicides over a long period of time exerts strong selection pressures on weed populations resulting in the evolution of resistance (Gaines et al., 2020; Powles and Yu, 2010). The evolution of resistance to herbicides in weeds is analogous to the evolution of antibiotic resistance in certain strains of bacteria in hospitals (Haywood et al., 2021). A weed plant is defined as herbicide resistant when it has inherited the ability to survive and reproduce when exposed to a dose of herbicide which is usually lethal (WSSA, 1998). This is distinct from herbicide tolerance which is the inherent ability of a plant species to survive and reproduce after herbicide treatment – implying no selection or genetic manipulation has occurred to make the plant tolerant (WSSA, 1998). The evolution of herbicide resistance in weeds was first predicted by Harper (1956), and the first case was reported soon after when a population of *Commelina diffusa* growing among sugarcane in Hawaii was found to have

evolved resistance to applications of the auxin mimicking herbicide 2,4-D (Hilton, 1957). Continued intense use of a finite number of herbicide MOA since then has resulted in a large increase in the number of herbicide resistance cases over time (Fig. 1.1). The persistent use of a limited set of herbicides has driven the evolution of herbicide resistance as weeds adapt to survive and reproduce in the face of prolonged selective pressure to a single or few herbicides. There are now 522 reported unique cases of herbicide-resistant weeds globally, with 269 weed species evolving resistance to 21 of the 26 herbicide MOA classes (Heap, 2023).

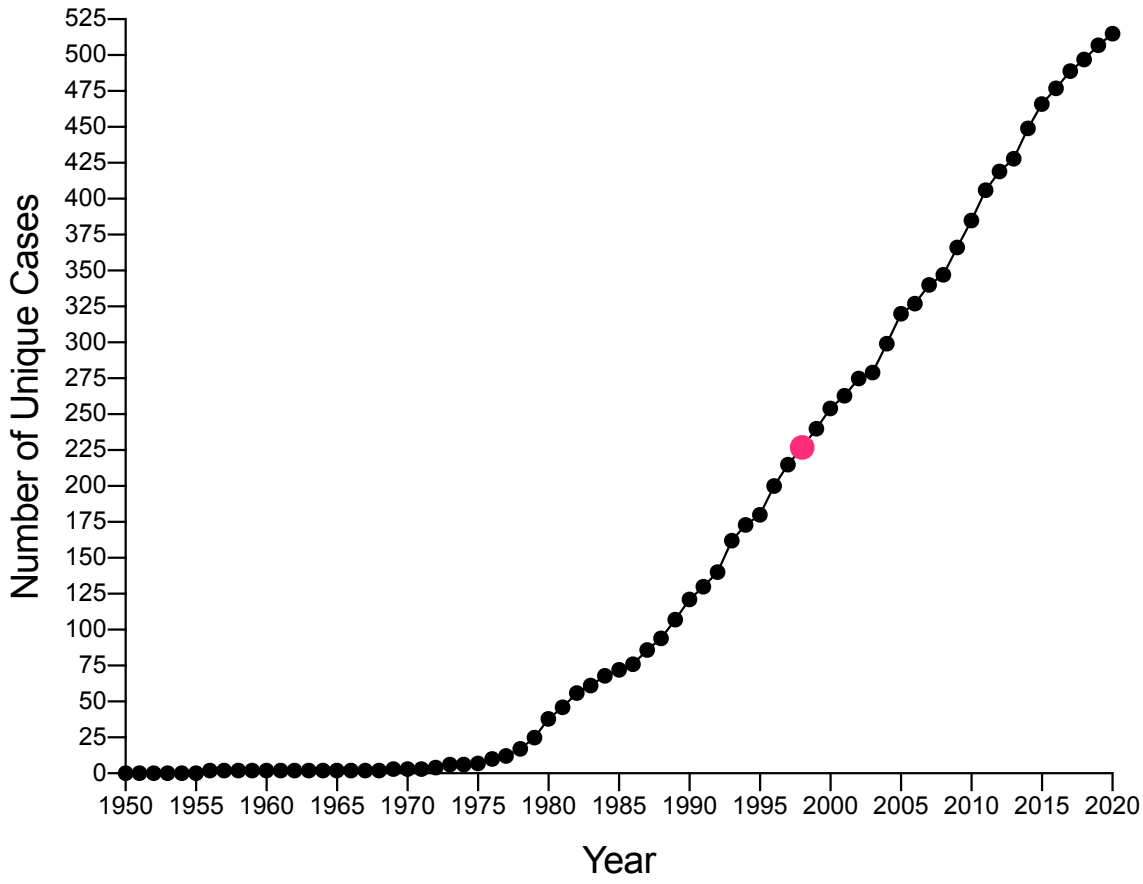


Fig. 1.1 Herbicide resistance cases have increased since the widespread adoption of herbicides
A unique case of herbicide resistance is a weed species with resistance to a herbicide of a particular site of action. The year the first case of glyphosate resistance was reported is highlighted in pink. Redrawn from (Heap, 2023).

Until the 1990s, the evolution of herbicide resistance was not a threat to food production, because farmers could rely on new herbicide MOAs being introduced to the market (Duke, 2012). This was considered the 'golden era' of herbicide discovery, in which a new herbicide MOA was brought to market approximately every 2 years between the 1950s and mid-1980s (Dayan, 2019). The availability of novel herbicides to farmers meant farmers were usually able to deploy new herbicides to combat weed populations that had evolved herbicide resistance. However, in the 30 years following the 1990s, the rate at which novel herbicides were brought to market rapidly declined (Gerwick, 2010). This meant that farmers repeatedly and extensively used herbicides with few MOAs, which resulted in the evolution of widespread resistance. This pattern of overreliance on a limited set of herbicides accelerated the evolution of resistance in numerous weed species. This usually occurred in response to application of the most used herbicides such as atrazine, glyphosate, and tribenuron-methyl (Heap, 2023). This shift from frequent herbicide discovery before the mid 1980s to almost no new discoveries after the early 1990s resulted in the overuse of a limited range of herbicides, driving rapid resistance evolution in many weed species.

There are a variety of reasons for the reduction in the rate of herbicide discovery (Duke, 2012). These include, an increased cost of herbicide development, regulatory barriers to herbicide commercialisation and the development and widespread adoption of glyphosate resistant (GR) crop technologies in the mid- to late-1990s. The price of glyphosate also dropped considerably following the expiration of its patent, exacerbating the situation (Gianessi, 2008). These factors have subsequently led to an economically unfavourable market for companies involved in herbicide discovery, which has led to a widespread reduction in the number of companies searching for novel herbicides. The impact of economic pressure on the number of companies involved in herbicide discovery is highlighted in Fig. 1.2.

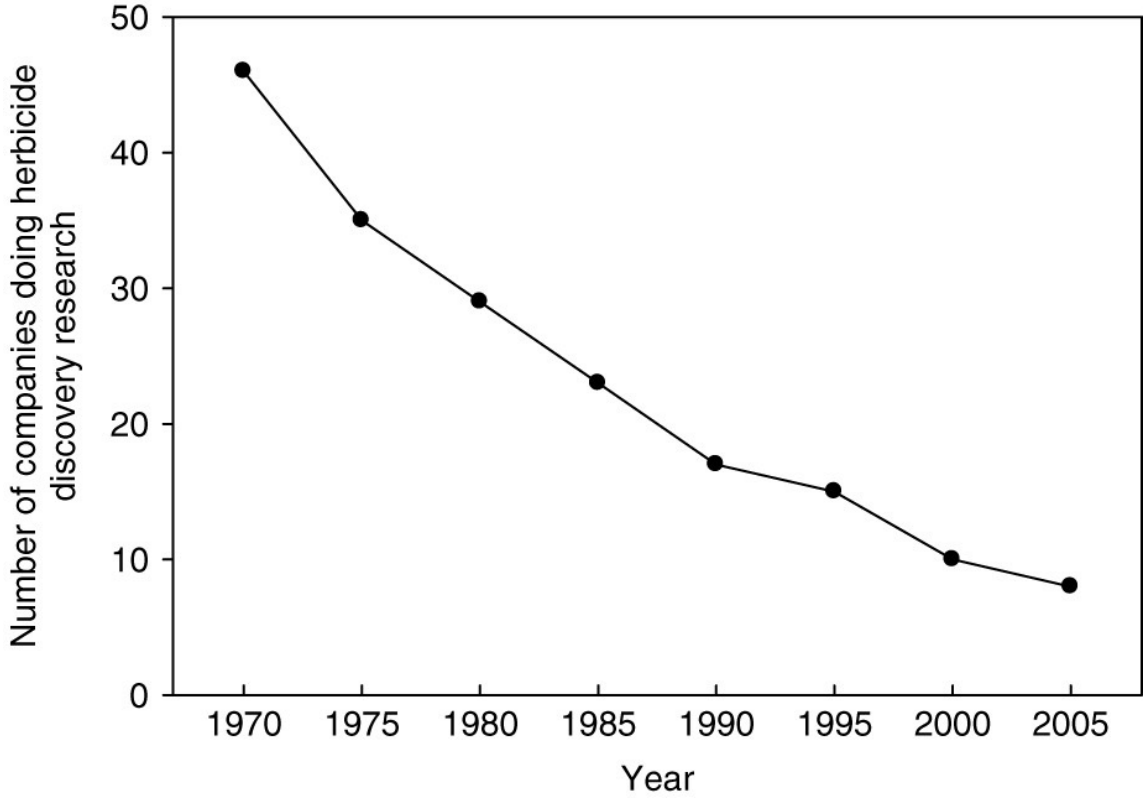


Fig. 1.2 Chronological decrease in the number of companies searching for new herbicides The number of companies involved in herbicide discovery has continuingly decreased in response to economic pressure caused by, at least in part, regulatory barriers to herbicide commercialisation and the success of one herbicide – glyphosate. Chart sourced from (Duke, 2012).

The lack of herbicide discovery and overuse of the few MOAs available to farmers has driven widespread evolution of herbicide resistance in populations of weeds. This has now been recognised as an urgent problem which threatens food security (Hawkins et al., 2019). As a result, extensive research has been carried out on the mechanisms through which weeds have evolved resistance to herbicides.

1.4. Mechanisms of herbicide resistance

Weeds have evolved resistance to herbicides through a variety of mechanisms. Understanding the genetic and mechanistic bases of herbicide resistance evolution gives opportunities to overcome the problem by fuelling technological innovation that can restore efficient, cost-effective, and sustainable weed control in the future. Furthermore, certain resistance mechanisms may be utilised in transgenic crops to improve the efficacy of current herbicides by increasing their selectivity. Herbicide resistance mechanisms can be broadly

separated into two categories, target-site resistance (TSR) and non-target site resistance (NTSR) (Délye et al., 2013). I will briefly describe each group of mechanism here:

1.4.1. Target-site resistance (TSR)

Generally, herbicides function by binding to a target protein, catastrophically inhibiting the protein's function, leading to plant death (Gaines et al., 2020). TSR occurs because mutations cause alterations in the target of the herbicide which result in a reduction of the inhibition of the target protein by the herbicide (Devine and Shukla, 2000; Murphy and Tranel, 2019). Often, TSR arises from point mutations in the gene encoding the target of the herbicide, leading to a change in the amino acid sequence of the target protein, and therefore a structural change which alters the herbicide's affinity to bind to its target. Since most herbicide targets are encoded by a single gene, TSR is generally monogenic and usually simple (Délye et al., 2013; Yuan et al., 2007). TSR can also be conferred by overexpression of the herbicide target by gene amplification or alterations in a gene promoter (Powles and Yu, 2010). The overexpression of the herbicide target results in herbicide resistance because, although the herbicide can still bind to its target, there remains a sufficient concentration of unbound, active target that allows the plant to survive.

1.4.2. Non-target site resistance (NTSR)

Non-target site resistance mechanisms are those which confer resistance by decreasing the amount of the active herbicide which reaches the target in the plant or those which alleviate the herbicide's toxic effects (Powles and Yu, 2010). Often, NTSR mechanisms are a subset of plant responses to abiotic stresses which have been altered in some way to instead give resistance to a herbicide (Délye et al., 2013). NTSR can also comprise physical mechanisms on the whole-plant level. For example, plants may evolve resistance through decreased uptake or translocation of a herbicide (Powles and Yu, 2010). On a cellular level, plants can detoxify herbicides, converting the active herbicide to a less toxic product. Herbicides can also be sequestered into sub-cellular compartments, such as the vacuole, where they are unable to interact with the target protein. These mechanisms can occur independently, simultaneously, or sequentially (Yuan et al., 2007). Herbicide detoxification is often carried out by members of large protein families such as cytochrome P450 monooxygenases (P450s), glutathione S-transferases (GSTs), glycosyltransferases (GTs), and ATP-binding

cassette (ABC) transporters (Gaines et al., 2020; Yuan et al., 2007). In the field, NTSR often evolves as a result of mutations which cause the overexpression of such proteins (Franco-Ortega et al., 2021; Leslie and Baucom, 2014). Overexpression of proteins from these large families can confer herbicide resistance in transgenic plants in the lab (Mentewab and Stewart, 2005; Saika et al., 2014; Windsor et al., 2003). Unique to NTSR, this form of resistance can sometimes give plants the ability to resist the application of multiple herbicides with independent MOAs. This phenomenon is known as multiple-herbicide resistance and is especially problematic as it further reduces the number of effective herbicides against such weeds (Peterson et al., 2018).

1.5. NTSR: current knowledge

NTSR can encompass a variety of different mechanisms of evolved herbicide resistance. A common mechanism which has been reported in the literature is herbicide detoxification. The detoxification of herbicides can be grouped into four distinct phases: phase I conversion, phase II conjugation of the compound, phase III compartmentation and phase IV which is any further detoxification of the herbicide occurring in the vacuole or a cellular space following compartmentation (Hatzios et al., 1997). Members of four large enzyme families are usually responsible for each these phases of detoxification (Yuan et al., 2007). Phase I detoxification is typically carried out by P450 enzymes or other mixed function oxidases/peroxidases (Schuler, 1996; Siminszky, 2006; Yun et al., 2005). Reactions catalysed by these enzymes increase the polarity of the herbicide and can make the herbicide a substrate for conjugation to a secondary molecule such as glutathione or glucose (Morant et al., 2003). Such conjugation reactions comprise phase II detoxification and are often carried out by proteins such as GSTs or uridine diphosphate (UDP) glycosyltransferases (UGTs). Phase II detoxification involves the conjugation of a hydrophilic molecule to a polarised herbicide molecule using thiols or sugars (Jones and Vogt, 2001; Reade et al., 2004). A conjugated herbicide may then be more readily recognised by enzymes involved in compartmentation (phase III), namely ABC transporters (Conte and Lloyd, 2011). A summary of these phases of herbicide detoxification can be found in Table 1.1. The next section of my introduction will comprise a review of the current literature on the four main enzyme

families involved in herbicide detoxification, as well as a brief overview of other NTSR mechanisms such as reduced translocation and rapid cell death.

	Process	Solubility	Translocation	Phytotoxicity	Bioavailability	Example(s)
Initial Properties	Parent Compound	Lipophilic	Selective mobility	Toxic	Available	
Phase I	Oxidation Reduction Hydrolysis	Amphophilic	Modified or reduced mobility	Modified or less toxic	Less available	The P450 CYP749A16 in cotton (*Gossypium hirsutum*) confers tolerance to trifloxysulfuron sodium (Thyssen et al., 2018)
Phase II	Conjugation	Hydrophilic	Limited or immobile	Greatly reduced or non-toxic	Less available	GSTU4/19 from maize and GSU5, GSTU9 and GSTU10 from soybean (*Glycine max*) showed GSH conjugating activity towards chlorimuron ethyl (McGonigle et al., 2000). Two UGTs (GT83A1 and GT75D1) isolated in NTSR shortawn foxtail (*Alopecurus aequalis)* resistant to mesosulfuron-methyl (Zhao et al., 2017).
Phase III	Compartmentation	Hydrophilic and insoluble	Immobile	Non-toxic	Unavailable	ABC transporter EcABCC8 up-regulated in resistant *Echinochloa colona* and endows glyphosate resistance in transgenic rice (Pan et al., 2021)

Table 1.1 Herbicide detoxification can be separated into 3 distinct phases

1.5.1. Cytochrome P450s (P450s)

Cytochrome P450s (P450s) are the largest family of enzymes in plants and carry out a variety of functions (Xu et al., 2015). P450s predominantly catalyse reactions in which electrons from NADPH are used to cleave atmospheric dioxygen, a process facilitated by the enzyme NADPH-CYP reductase. Subsequently, one oxygen atom undergoes hydroxylation with the substrate, while the second oxygen atom combines with hydrogen to form water. These reactions produce a more reactive and soluble product molecule. This function detoxifies

xenobiotic molecules. Plant P450s are more numerous and diverse than in metazoans, which may be an adaptation to plant's sessile life on land (Coleman, 1997). In eukaryotes, P450s are membrane-bound and are usually localised in the endoplasmic reticulum. In addition to their known xenobiotic detoxification role in plants, P450s also have a role in drug metabolism in humans and pesticide resistance in insects (Li et al., 2007; Danielson, 2002).

P450 enzymes are frequently associated with phase I of herbicide detoxification (Dimaano and Iwakami, 2021). Evidence supporting the role of P450s in herbicide detoxification was first reported to the photosystem II (PSII) inhibiting herbicide monuron in cotton seedlings (Frear et al., 1969). Since then, there has been an accumulation of evidence for the involvement of P450s in herbicide metabolism and therefore herbicide detoxification in both weed and crop species (Burnet et al., 1993; Cole and Owen, 1987; Fonné-Pfister et al., 1990; Kreuz and Fonné-Pfister, 1992; Mougin et al., 1991; Ohkawa et al., 1999; Powles and Yu, 2010; Siminszky, 2006). P450s play a role in the metabolism of herbicides from various classes and in particular the ALS and PSII inhibitors (Fonné-Pfister et al., 1990; Siminszky, 2006; Xiang et al., 2005). Evidence has also been reported that P450s play a role in the multiple herbicide resistance which has evolved in some weeds (Busi et al., 2011; Iwakami et al., 2014). For example, it has been reported that the overexpression of a single P450 gene in transgenic plants can confer resistance to up to 13 different herbicides (Hirose et al., 2005). In summary, P450 enzymes have been identified as key contributors to herbicide detoxification of multiple classes of herbicide, while also being reported to play a role in instances of multiple herbicide resistance.

1.5.2. Glutathione S-transferases (GSTs)

Despite being reported less frequently in the literature than P450s, a large body of evidence has also accumulated for the role of glutathione S-transferases (GSTs) in NTSR (Nandula et al., 2019). GSTs contribute to herbicide detoxification by catalysing the conjugation of glutathione, which is hydrophilic, to a hydrophobic herbicide molecule. This usually has the effect of reducing the toxicity of the herbicide, its mobility within the plant and its bioavailability (Hatzios et al., 1997). These conjugates may then be transported from cells or into the vacuole by membrane-bound transporter proteins (Cole and Deeley, 2006). GSTs are a large and diverse enzyme family, and this diversity allows them to conjugate a large range

of molecules with glutathione in order to detoxify them (Cummins et al., 2011; Yuan et al., 2007). GSTs are involved in the detoxification of various herbicide classes including PSII inhibitors, very long-chain fatty acid biosynthesis inhibitors, and ACCase inhibitors (Cummins et al., 2013; Karavangeli et al., 2005; Li et al., 2017; Thom et al., 2002; Shah et al., 1986). *A. myosuroides* (blackgrass) is a weed species which is causing many problems in western Europe due to its evolution of high levels of resistance to multiple herbicide MOAs. It has been reported that GSTs may play a significant role in contributing to herbicide resistance in NTSR populations of this weed, including by acting as glutathione peroxidases to confer resistance to multiple herbicides (Cummins et al., 1999, 1997; Letouzé and Gasquez, 2003).

1.5.3. Uridine diphosphate (UDP) glycosyltransferases (UGTs)

UDP-glycosyltransferases (UGTs), which make up family 1 of the glycosyltransferase enzyme superfamily, are another group of enzymes which are also involved in phase II (conjugation) of herbicide detoxification. UGTs catalyse the transfer of glycosyl residues from activated nucleotide sugars, such as uridine diphosphate glucose, to acceptor molecules (Yonekura-Sakakibara and Hanada, 2011). The addition of a glycosyl residue to a herbicide usually makes it more hydrophilic and contributes to herbicide detoxification. The conjugation of herbicides with glucose mediated by UGTs was first reported in studies which treated chlorsulfuron-tolerant wheat with C^{14}-labelled chlorsulfuron and showed that following phenyl ring-hydroxylation by P450s, the chlorsulfuron was rapidly conjugated with glucose (Sweetser et al., 1982). There are fewer studies reporting the role of UGTs in herbicide detoxification than for P450s and GSTs.

UGTs contribute to NTSR in multiple herbicide-resistant populations of blackgrass (Brazier et al., 2002). Furthermore, a UGT was also consistently overexpressed in populations of *Lolium rigidum* (rigid ryegrass), which had evolved NTSR to diclofop; an ACCase inhibitor (Gaines et al., 2014). Rigid ryegrass is a problematic weed in Australia which has evolved resistance to many herbicide MOAs. More recently, a genome-wide association study (GWAS) of populations of glyphosate-resistant and glyphosate sensitive *Amaranthus tuberculatus* in North America showed that many UGTs had increased copy number in resistant populations, and therefore likely played a role in contributing to glyphosate resistance in these

populations (Kreiner et al., 2021, 2019). In summary, UGTs have been widely associated with NTSR in various herbicide-resistant weed populations, highlighting their significant role in resistance evolution.

1.5.4. ATP binding cassette (ABC) transporters

Whereas P450s, GSTs, and UGTs detoxify herbicides through biochemical modification, proteins from the ABC transporter family can confer herbicide resistance by compartmentation of herbicides, or their toxic metabolites, as part of phase III detoxification (Yuan et al., 2007). ABC transporters use energy from ATP hydrolysis to drive active transport of substances such as nutrients, secondary metabolites, phytohormones, proteins and peptides, and xenobiotics across membranes (Kang et al., 2011). ABC transporters are localised to membranes and in vascular plants have been characterised as having a wide range of functions such as the excretion of toxic compounds, the translocation of fatty acids, and the transport of heavy metals, and the general maintenance of cell homeostasis (Schulz and Kolukisaoglu, 2006). Like P450s, ABC transporters are one of the most diverse gene families in plants and this diversity of genes in plants is far higher than is found in non-plant species (Henikoff et al., 1997). This diversity may play a key role in the inherent ability of these proteins to confer resistance to molecules, such as herbicides, that plants have only been exposed to very recently in their evolutionary history.

Herbicide metabolites have often been reported to accumulate in plant vacuoles and this consistent with transport by ABC transporter proteins. However, there is little evidence which associates this process with evolved herbicide resistance in weeds. Herbicide tolerance has, however, been directly associated with ABC transporters in various model species and crop plants (Yuan et al., 2007). For example, populations of *Echinochloa colona*, evolved NTSR to glyphosate in western Australia. An ABC transporter, EcABCC8, is up regulated in resistant plants but a direct link between ABC transporter upregulation and herbicide resistance was not in this species (Pan et al., 2021). However, the gene encoding this protein was then inserted in to rice and *ABCC8* transgenic rice plants were resistant to glyphosate. It was reported that ABCC8 likely serves to export glyphosate to the apoplast, reducing the levels of glyphosate interacting with its target in the chloroplast.

1.5.5. Reduced translocation and rapid cell death

Many foliar applied herbicides rely on translocation through the phloem to reach their target and have optimal activity (Gaines et al., 2020). Weeds can evolve resistance to these herbicides by limiting their translocation, retaining them in the leaves to which they were applied and restricting their movement to actively growing parts of the plant. NTSR to glyphosate has been reported to evolve this way in rigid ryegrass (Gaines et al., 2019). Glyphosate was found to accumulate at the edges of leaves in glyphosate-resistant rigid ryegrass, resulting in less of the herbicide reaching the meristems than in glyphosate sensitive rigid ryegrass (Lorraine-Colwill et al., 2002). A similar mechanism that restricts herbicide mobility confers resistance to paraquat in the model plant species *Arabidopsis thaliana*. Characterisation of a mutant paraquat-resistant *A. thaliana* line revealed the cause was the silencing of a single gene, Paraquat Resistance 1 (*PAR1*), which is a putative amino acid transporter. A loss-of-function of this protein is thought to confer resistance to paraquat due to reduced active transport of paraquat into the chloroplast where it is active (Li et al., 2013). Necrosis and rapid cell death-based resistance has also been reported to have evolved against 2,4-D and confer resistance to this herbicide in the weed *Conyza sumatrensis*, however the genetic mechanisms underlying this resistance case is still unknown (Queiroz et al., 2019). Such mechanisms may be an example of plants utilising existing disease or general stress responses for herbicide resistance.

1.5.6. NTSR summary

Sustained applications of herbicides to weeds have led to the evolution of herbicide resistance through many mechanisms. NTSR can occur through herbicide detoxification or other mechanisms which ultimately reduce the amount of active herbicide reaching its target. Enzymes from four families have most often been linked to detoxification processes and resistance often evolves due to increased expression of genes encoding enzymes from these four families as a result of gain-of-function mutations. To date, very little work has been done to resolve the evolutionary history of these four enzyme families in plants. Understanding the evolutionary history and function of these protein families may help us to understand how they can evolve resistance to herbicides in the field. A phylogenetic analysis

of two of these protein families, UGTs and ABC transporters, comprises Chapter 2 of my thesis. Emerging evidence suggests that there are many other mechanisms of NTSR resistance in weeds. Identifying which genes can confer herbicide resistance in plants is paramount to finding a solution to the herbicide resistance problem. This is another major focus of my thesis.

1.6. Glyphosate and glyphosate resistance

Glyphosate [*N*-(phosphonomethyl)glycine] has a unique mode of action among herbicides by inhibiting the enzyme 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS), which catalyses a key step of the shikimate pathway (Fig. 1.3). Therefore, glyphosate is thought to kill plants by disrupting the biosynthesis of the aromatic amino acids tryptophan, phenylalanine, and tyrosine. EPSPS catalyses the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) to shikimate-3-phosphate (S3P) to form 5-enolpyruvylshikimate 3-phosphate (EPSPS) and inorganic phosphate (Bentley et al., 1990; Levin and Sprinson, 1964). Glyphosate inhibits EPSPS by binding almost irreversibly to the active site and does so competitively in regard to PEP (Schönbrunn et al., 2001). Inhibition of EPSPS by glyphosate leads to the approximately 20% of photosynthetically fixed carbon flowing through the shikimate pathway being stalled and converted to shikimate (Duke, 1988; Herrmann, 1995). Although the primary hypothesis of how glyphosate kills plants is starvation of aromatic amino acids, it has also been hypothesised that increased carbon input to the shikimate pathway due to dysregulation as a result of EPSPS inhibition could lead to carbon shortages in other pathways and therefore plant death (Siehl, 1997). I explore these potential causes of plant death as a result of glyphosate treatment in Chapter 5 of my thesis.

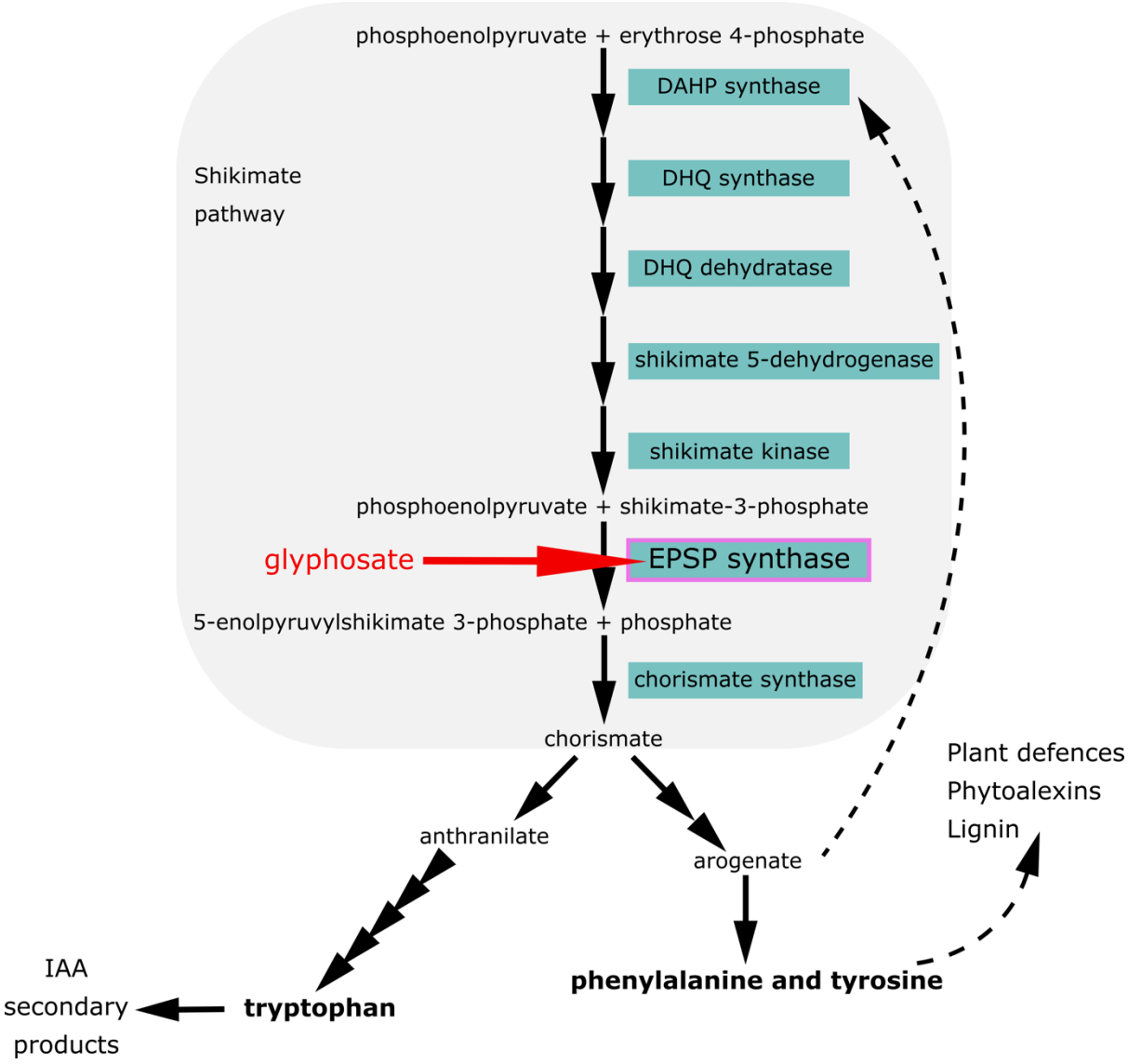


Fig. 1.3. Shikimate pathway in plants
Scheme of the shikimate pathway in plants with the site of inhibition by glyphosate highlighted. Products of the pathway and feedbacks (dashed arrows) are shown. Adapted from (Duke and Powles, 2008).

Today, glyphosate is the most dominant herbicide in the world in terms of both amount used and capital market value, which in 2022 was estimated at USD 9.6 billion ("www.verifiedmarketresearch.com/product/glyphosate-market," 2023). Monsanto introduced Glyphosate to the market in 1974, under the brand name RoundUp. Initially, its usage remained comparable to that of other foliar applied herbicides, as its lack of selectivity limited its widespread adoption. However, introduction of glyphosate-resistant (GR) crops changed this. GR technology increased the selectivity of glyphosate and therefore allowed it to be used post-emergence. GR soybean (*Glycine max*) was the first GR crop and

was brought to the market in the USA in 1996 (Duke, 2005). This transgenic crop utilised a resistant form of the EPSPS enzyme from a glyphosate-resistant strain of *Agrobacterium tumefaciens* (CP4) (Padgette et al., 1995). By 2010, seven GR crops were in use commercially: soybean, canola (oilseed rape) (*Brassica napus*), Polish canola (*Brassica rapus*), alfalfa (*Medicago sativa*), and sugar beet (*Beta vulgaris*) (Green, 2009). The rapid adoption of GR crops (in the Americas at least), and the expiration of glyphosate patents in 2000, which greatly reduced the cost of the herbicide itself, meant that the use of glyphosate increased rapidly. This rapid increase in use was correlated with an increase in the number of weed populations evolving resistance to this herbicide following the first reported case of evolved glyphosate resistance in 1998 (Powles et al., 1998; Pratley et al., 1999; Duke and Powles., 2008).

The continued intense selection pressure exerted on weeds and crops by glyphosate application means that there are now more cases of resistance to glyphosate than all but one other herbicide (atrazine) (Fig. 1.4) (Heap, 2023). Due to glyphosate's importance in agricultural, it is paramount we understand glyphosate resistance mechanisms in the hope that we can overcome them and continue to efficiently use this herbicide in the future.

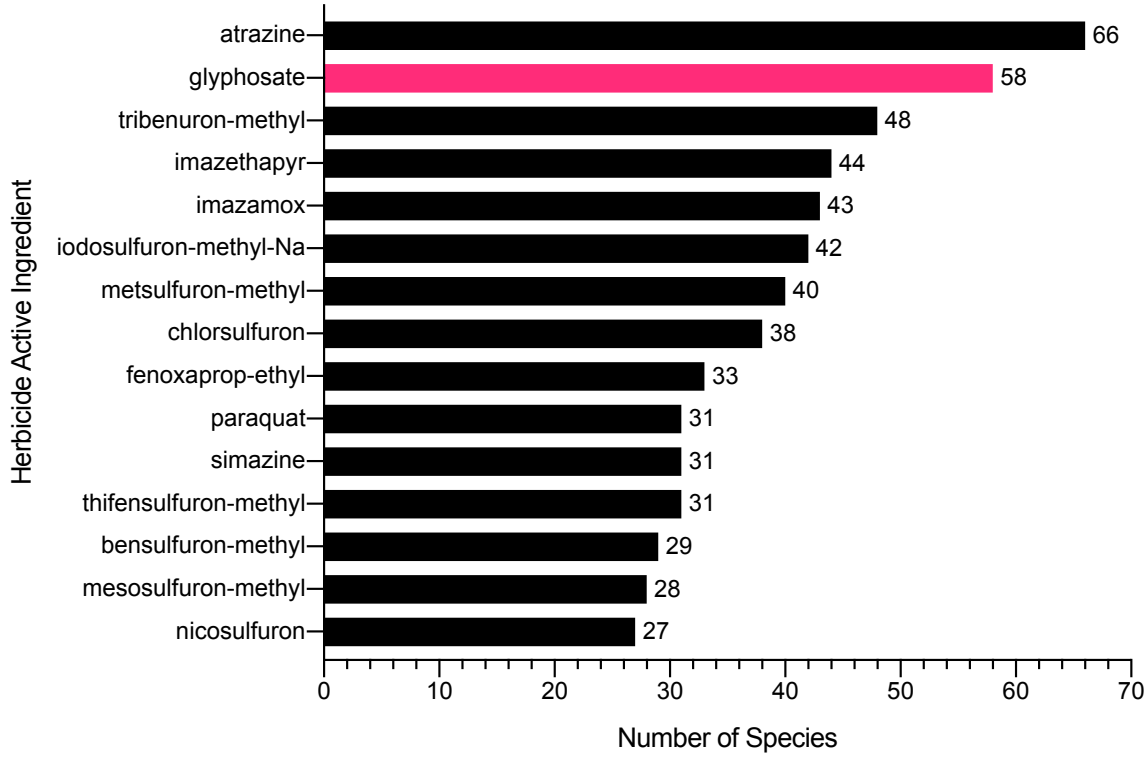


Fig. 1.4. Number of resistant weed species to 15 active herbicide ingredients
Glyphosate has the second highest number of resistance cases of any herbicide. Number of resistance cases for each of the herbicides with the top 15 number of weed species with resistance against them is plotted. Redrawn from (Heap, 2023).

There are currently 356 reported cases of glyphosate resistance (Heap, 2023), and weeds have evolved resistance to glyphosate through a variety of mechanisms. TSR has evolved in weeds to glyphosate almost exclusively through mutations which lead to a change in the Pro-106 residue of the enzyme (Gaines et al., 2020; Sammons and Gaines, 2014). These mutations lead to substitution of the proline for serine, threonine, alanine, or leucine (Heap, 2023). Such mutations cause a change in the structure of the EPSPS active site which reduce the enzymes affinity for glyphosate, but also its usual substrate PEP (Healy-Fried et al., 2007). Another mutation in a similar region can overcome this loss of affinity for PEP, however. Such a mutation has been commercialised in the GA21 GR maize (Ravinder S. Sidhu et al., 2000). In this GR maize the glyphosate-resistance endowing P106S mutation is accompanied a T102I mutation which leads to higher affinity to PEP than the P106S alone, but a maintained low affinity to glyphosate (Funke et al., 2009). This double T102I, P106S mutation is known as TIPS. This mutation has now also been identified to evolve in the field. It was first identified in populations of goosegrass (*Eleusine indica*) (Han et al., 2017), and

has also been reported in a population of hairy beggarticks (*Bidens pilosa*) (Alcántara-de la Cruz et al., 2016). More recently, a weed species in which the stepwise evolution of a triple amino acid substitution in its EPSPS sequence was identified in Argentina (Perotti et al., 2019). This was in smooth pigweed (*Amaranthus hybridus*) and consisted of T102I, A103V and P106S mutations, which was found to confer very high levels of resistance to glyphosate (García et al., 2019). Such stepwise evolution of resistance mutations is indicative of continued use of glyphosate even in the face of resistance evolution – a symptom of the few herbicide options available to farmers.

TSR to glyphosate can also be conferred by overexpression of the glyphosate target gene, *EPSPS*. This mechanism of resistance has evolved many times in the field. For example, some glyphosate-resistant populations of *E. indica* in China overexpress *EPSPS* up to 29.7 times that of sensitive populations, leading to an almost 10-fold increase in glyphosate resistance (Chen et al., 2020). Furthermore, some of these populations overexpressing *EPSPS*, were overexpressing a form encoding a P106A substitution in the EPSPS enzyme. Overexpression of a mutated form of *EPSPS* had been reported previously in *E. indica* (Gherekhloo et al., 2017). Such examples highlight the persistence of resistance evolution when application of a single herbicide is continued even after resistance evolves. *EPSPS* overexpression has also been identified to contribute to glyphosate resistance in a variety of other weeds, such as *A. palmeri, Salsola tragus*, *L. rigidum* and *C. banariensis* (Baerson et al., 2002; Gaines et al., 2020b, 2010; Yanniccari et al., 2023). Many molecular mechanisms lead to *EPSPS* overexpression. One extraordinary example was the formation of an approximately 400 kbp long extrachromosomal circular DNA eccDNA replicon in the genome of *A. palmeri*, which led to the overexpression of *EPSPS*, along with a suit of other genes (Molin et al., 2020). Since overexpression of *EPSPS* can come about due to a variety of genetic and non-genetic changes, it therefore can sometimes represent a more rapid evolutionary response to glyphosate application than nucleotide-base genetic mutations.

A variety of different NTSR mechanisms have been found to evolve in weed populations to glyphosate. Detoxification mechanisms of glyphosate resistance in plants have rarely been found, although there is an example of an aldo-keto reductase enzyme degrading glyphosate in *E. colona* (Pan et al., 2019). However, since glyphosate is a foliar applied herbicide, NTSR

can evolve through mechanisms which prevent it translocating from the leaves to other growing parts of plants (Menendez et al., 2014). One dramatic example has evolved in giant ragweed (*Ambrosia trifida*) where, despite the usual slow action of glyphosate, plants of this species have evolved a rapid cell death response to glyphosate treatment. Treated cells die rapidly, preventing translocation to the meristem from which the plant regrows (Moretti et al., 2018; Van Horn et al., 2018). This has come to be known as the 'phoenix phenomenon'. Many other mechanisms of reduced translocation do not involve death of large parts of the plant. For example, several populations of *L. rigidum* in Australia are resistant to glyphosate due to a reduction in the amount of glyphosate translocating to the meristems compared to in sensitive populations (Wakelin et al., 2004). A similar mechanism was also found to be present in *Conyza canadensis*, in which glyphosate was still able to enter the symplasm of treated leaves but was not translocated to the meristems due to rapid sequestration into the vacuole (Ge et al., 2010). The authors describing this mechanism suggest that a transport protein such as an ABC transporter protein may be sequestering glyphosate into the vacuole in *C. canadensis*. Glyphosate resistance in *C. canadensis* was later found to be correlated with the expression levels of two ABC transporter genes (Nol et al., 2012; Tani et al., 2016). An ABC transporter, ABCC8, has also been found to confer glyphosate resistance in *E. colona* (Pan et al., 2021). Other mechanisms of transport have been suggested to confer glyphosate resistance, including those which limit the transport of glyphosate into the chloroplast (Sammons and Gaines, 2014). Novel mechanisms of NTSR to glyphosate are continually being discovered, and often these are not the most well-known mechanisms of herbicide detoxification. It is therefore important to generate systems in which we can research and predict potentially novel mechanisms of glyphosate resistance.

1.7. Using *Marchantia polymorpha* to study glyphosate resistance

Most research on herbicide resistance, has to date, been carried out in weed and crop species. However, understanding resistance mechanisms in these species is often challenging due to complex genetics and difficulties propagating weed and crop species under laboratory conditions. I instead propose to primarily study glyphosate resistance mechanisms in the model organism *Marchantia polymorpha*. There is anecdotal evidence that *M. polymorpha* is naturally glyphosate tolerant (Marble et al., 2017). It therefore

represents a unique opportunity among land plants to study non-evolved resistance mechanisms to glyphosate, which may be novel. In my thesis, I will also be using the model plant species *A. thaliana* and *Physcomitrium* patens and the wild plant species *Conocephalum conicum* and *Lunularia cruciata* to compliment the work done with *M. polymorpha*.

The suspected natural glyphosate tolerance in *M. polymorpha* is not the only asset which makes this species valuable for glyphosate resistance research. *M. polymorpha* is a dioecious liverwort species which is widespread across temperate regions of the world and has been used as a model system for many years (Shimamura, 2016). It is advantageous as a model system in general, and as a system to study glyphosate resistance, as it has a short generation time (3 months with sexual reproduction, 2-3 weeks with asexual reproduction), proves easy to propagate and cross, has a haploid dominant life cycle and has a small genome size of approximately 280 Mb (Bowman et al., 2017). The dominant haploid life cycle of *M. polymorpha* is of particular advantage in generating loss-of-function mutants, by for example CRISPR/Cas9 mutagenesis, as phenotypes caused by the mutations are visible in the first generation of transformed plants. This bypasses the lengthy generation of homozygous mutant lines necessary in other model species such as *A. thaliana*. In addition to these advantages, another advantage of *M. polymorpha*, especially for herbicide resistance research, is that it produces clonal propagules called gemmae. These provide a quick and easy system to quantify herbicide inhibition by growing the propagules on media supplemented with a herbicide and imaging from above. This can allow efficient comparisons of the effect different mutations and herbicide treatment regimes on herbicide resistance.

A potential drawback of using *M. polymorpha* to study glyphosate resistance mechanisms is that it is distantly related to the angiosperm weeds which are primarily targeted with glyphosate (Carpenter et al., 2019; Harris et al., 2020; Morris et al., 2018). This may limit the transferability of conclusions made relating to glyphosate resistance in *M. polymorpha* to angiosperms weeds. However, due to its technical advantages over angiosperm plants, *M. polymorpha* likely still represents a valuable tool in which to rapidly research glyphosate resistance mechanisms despite its evolutionary distance to angiosperm weeds.

The land plants (embryophytes) are a diverse taxon comprising more than 370,000 extant species which form a monophyletic group within the streptophyte lineage (Christenhusz and Byng, 2016; de Vries and Archibald, 2018). *M. polymorpha* is a member of the bryophytes which are widely considered to be a monophyletic clade consisting of the liverworts, mosses, and hornworts (Harris et al., 2022, 20; Nishiyama et al., 2004; Sousa et al., 2019; Su et al., 2021). The embryophytes comprise of two monophyletic clades; bryophytes and the vascular plants (tracheophytes) (Donoghue et al., 2021; Harris et al., 2022). *M. polymorpha* (bryophyte) and angiosperm weeds (tracheophytes) therefore represent distinct clades of the embryophyte lineage. Latest estimates predict the bryophytes and tracheophytes to have last shared a common ancestor as early as during the Cambrian Period (515 mya) and as late as the early Devonian Period approximately 415 mya (reviewed in Donoghue et al., 2021; Harris et al., 2022). Following the divergence of these embryophyte clades, the bryophyte lineage lost many genes associated with the development of vasculature and stomata (Harris et al., 2020, 2022). However, many of the gene networks in the gametophyte generation of *M. polymorpha* are similar to those present in the dominant sporophyte generation of angiosperms (Bowman et al., 2017; Ligrone et al., 2012; Pires and Dolan, 2012). This could facilitate the application of findings made regarding the mechanisms of resistance to glyphosate and response to glyphosate treatment in *M. polymorpha,* to angiosperm weeds.

One fundamental difference between *M. polymorpha* and angiosperms, which may influence its response to glyphosate, is the structural properties of the chloroplast wall in *M. polymorpha*. The target of glyphosate, EPSPS, is chloroplast localised, so transit into the chloroplast is key to its inhibitory function. Chloroplasts are proposed to be derived from the endosymbiosis of a single cyanobacterial ancestor (Rodríguez-Ezpeleta et al., 2005; Whatley and Whatley, 1981). Peptidoglycan (murein) is a covalent macromolecular structure that is found in the cell walls of almost all eubacteria, and so it is likely the cyanobacterial ancestor of the chloroplast also had a cell wall containing peptidoglycan (Heijenoort, 2001). It was previously thought that peptidoglycan had been lost from the walls of the chloroplast. However, nine genes homologous to those necessary for peptidoglycan biosynthesis were identified in the *P. patens* genome (Machida et al., 2006). These genes are conserved in the

genome of *M. polymorpha* but many are absent from angiosperm genomes (Bowman et al., 2017). There is an increasing body of evidence that peptidoglycan is indeed synthesised in some green plants and that this accumulates around the chloroplast (Amanda J Dowson et al., 2022; Hofmann, 2016; Homi et al., 2009; Matsumoto et al., 2012). The conservation of peptidoglycan biosynthesis genes in *M. polymorpha* means that the presence of peptidoglycan around the chloroplast is likely also conserved in this liverwort. Consistent with this, application of the peptidoglycan biosynthesis antibiotic ampicillin to *M. polymorpha* disrupts chloroplast division (Tounou et al., 2002). Peptidoglycan may influence the movement of glyphosate into the chloroplast, however, since it is predicted glyphosate is actively transported across the chloroplast membrane, this is unlikely (Takano et al., 2019).

In this DPhil thesis, I investigate glyphosate resistance mechanisms in *M. polymorpha*. I conducted phylogenetic analyses of the UGT and ABC transporter enzyme families in 10 species of the Archaeplastida, finding that ABC transporters were present in the last common ancestor of the Archaeplastida, whereas UGTs are present only in the streptophyte lineage. I demonstrate that members of these families associated with herbicide resistance come from a diversity of clades. I then confirm and quantify the magnitude of glyphosate tolerance in *M. polymorpha* compared to the glyphosate-sensitive *A. thaliana*. This validates the use of *M. polymorpha* to study glyphosate resistance mechanisms. I then use purified plant and bacterial EPSPS enzyme assays and expressed At*EPSPS* and Mp*EPSPS* genes in transgenic *A. thaliana* plants to show that the *M. polymorpha* EPSPS is not glyphosate tolerant, and therefore this is not the cause of glyphosate tolerance in *M. polymorpha*. Using AlphaFold structural modelling, I then demonstrate that the peptidoglycan biosynthesis enzyme MpMurA, is structurally almost identical to MpEPSPS. I demonstrate that MpMurA contributes to glyphosate tolerance in *M. polymorpha* by genetically and chemically altering the function of this protein in this species. Then, by expressing Mp*MurA* in *A. thaliana*, I show that MurA contributes to glyphosate tolerance in *M. polymorpha* independently from peptidoglycan biosynthesis. Using purified enzyme assays, I then show that MpMurA catalyses the same reaction – that forms EPSP as a product – as MpEPSPS and conclude that this likely relates to the mechanism of glyphosate tolerance observed in *M. polymorpha*. I finally assess the metabolic response of glyphosate treatment in *M. polymorpha*, showing that this is drastically different to the response in *A. thaliana*.

To conclude, my thesis identifies the timings of the origin of important NTSR genes within the Archaeplastida, presents the kinetic parameters of specific plant EPSPS enzymes for the first time, and defines a novel mechanism of glyphosate tolerance involving a peptidoglycan biosynthesis enzyme.

Chapter 2: ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation

A version the following chapter was published in:
Caygill, S., Dolan, L., 2023. ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation. PLoS One 18, e0287356. https://doi.org/10.1371/journal.pone.0287356

Original PLoS One formatted paper in Appendix

Authors

Samuel Caygill[1,2], Liam Dolan[1,2]

Affiliations

[1]Gregor Mendel Institute, Vienna, Austria

[2]Department of Biology, University of Oxford, United Kingdom

Author Contributions

Writing – original draft: Samuel Caygill

Writing – review and editing: Liam Dolan

2.1 Abstract

ATP-binding cassette (ABC) transporters actively transport various substances across membranes, while uridine diphosphate (UDP) glycosyltransferases (UGTs) are proteins that catalyse the chemical modification of various organic compounds. Both of these protein superfamilies have been associated with conferring herbicide resistance in weeds. Little is known about the evolutionary history of these protein families in the Archaeplastida. To infer the evolutionary histories of these protein superfamilies, I compared protein sequences collected from 10 species which represent distinct lineages of the Archaeplastida – the lineage including glaucophyte algae, rhodophyte algae, chlorophyte algae and the streptophytes – and generated phylogenetic trees. I show that ABC transporters were present in the last common ancestor of the Archaeplastida which lived 1.6 billion years ago, and the major clades identified in extant plants were already present then. Conversely, UGTs were only identified in members of the streptophyte lineage, which suggests a loss of these proteins in earlier diverging Archaeplastida lineages or arrival of UGTs into a common ancestor of the streptophyte lineage through horizontal gene transfer from a non-Archaeplastida eukaryote lineage. Within the streptophyte lineage, most diversification of the UGT protein family was found to have occurred in the vascular lineage, with 17 of the 20 clades identified in extant plants present only in vascular plants. Based on my findings, I conclude that ABC transporters and UGTs are ancient protein families which diversified during Archaeplastida evolution, which may have evolved for developmental functions as plants began to occupy new environmental niches and are now being selected to confer resistance to a diverse range of herbicides in weeds.

2.2 Introduction

Populations of weeds are continually evolving resistance to herbicides (Heap, 2023). Herbicide resistance mechanisms can be classified into 2 general groups: target site resistance (TSR) and non-target site resistance (NTSR). Target site resistance (TSR) occurs as a result of mutations in the gene coding for the herbicide's target protein, which causes reduced inhibition of the herbicide target. Non-target site resistance (NTSR) can occur due to a variety of mechanisms, but ultimately results in a reduction in the amount of herbicide reaching its target protein, or by reducing damage caused by the herbicide to the plant (Powles and Yu, 2010). Some of the NTSR mechanisms reported include reduction in the uptake or translocation of the herbicide, metabolic detoxification of the herbicide and sequestration of the herbicide to a location where it is not active, such as the vacuole (Délye et al., 2015; Gaines et al., 2020; Ghanizadeh and Harrington, 2017; Yuan et al., 2007). Diverse genetic changes affecting NTSR genes can lead to their encoded proteins contributing to herbicide resistance. For example, mutations leading to gene overexpression, the expression of a hyperactive form of the protein, loss-of-function mutations or mutations leading to changes in substrate specificity, can all result in NTSR (Gaines et al., 2020), (Ito and Gray, 2006; Xi et al., 2012). The genetic basis of NTSR is poorly understood, but changes in the expression of both ATP-binding cassette (ABC) transporters and uridine diphosphate (UDP) glycosyltransferases (UGTs) have been shown to confer resistance to herbicides in weeds (Délye et al., 2015; Gaines et al., 2020; Yuan et al., 2007).

ABC transporters are one of the largest protein families in plants and are conserved across all domains of life (Henikoff et al., 1997). ABC transporter proteins use the energy of ATP hydrolysis to drive active transport of substances across membranes (Kang et al., 2011; Theodoulou, 2000). In plants, ABC transporters transport nutrients such as sugars, amino acids, and metal ions; secondary metabolites such as flavonoids, terpenoids and alkaloids; phytohormones such as auxins, cytokinins and gibberellins; proteins and peptides; and xenobiotics (Borghi et al., 2015; Martinoia et al., 2002; Theodoulou and Kerr, 2015). Plant ABC transporters are also involved in the translocation of toxic compounds, such as herbicides, into the vacuole and extracellular spaces, where they are no longer toxic (Ge et al., 2010; Martinoia et al., 1993). Numerous ABC transporters have been identified to

compartmentalise herbicides in this way, as part of a detoxification process that has evolved in some weeds with NTSR to at least one herbicide.

Most ABC transporters are membrane bound and consist of four subunits – two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs) (Supplementary Fig. S2.1) (Higgins, 1992). Generally, the substrate binds to the TMDs of the transporter on one side of the membrane. Binding of ATP to the NBDs then causes a conformational change in the transporter, closing the TMDs around the substrate. Subsequently, the hydrolysis of ATP causes another conformational change in the transporter that opens up the TMDs on the opposite side of the membrane, releasing the substrate (Kang et al., 2011; Locher, 2016). It is with this mechanism that ABC transporters are able to transport herbicides as part of a herbicide-resistance-conferring detoxification process.

Agricultural fields have been treated with herbicides since the 1940s, exposing plants to chemical diversity that they would not have previously encountered. Some ABC transporters catalyse the transport of a wide variety of synthetic compounds across membranes that they would not have encountered in their evolutionary history. ABC transporters are one of the most diverse protein families in plants, exhibiting significant variation in sequence, structure, and function. In plants, ABC transporters have been classified into 8 subfamilies (ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABCG and ABCI) (Kang et al., 2011; Saha et al., 2015; Theodoulou, 2000; Verrier et al., 2008). Previous phylogenetic analyses in plants have shown all ABC transporter subfamilies, excluding ABCI, constitute monophyletic clades (Kang et al., 2011; Theodoulou, 2000; Verrier et al., 2008). However, previous phylogenetic analyses of these protein families focused primarily on angiosperms. To gain a more comprehensive understanding of the evolution of this protein family, here I carry out a phylogenetic analysis of ABC transporters from across the Archaeplastida.

Uridine diphosphate (UDP) glycosyltransferases (UGTs) are part of the largest family (Family 1) of the glycosyltransferase superfamily, and this subset of UGTs have been widely associated with herbicide resistance (Dimunová et al., 2022; Yu et al., 2017). UGTs are present in bacteria, fungi, plants, and animals and catalyse the transfer of glycosyl residues

from activated nucleotide sugars, such as uridine diphosphate glucose, to acceptor molecules (Yonekura-Sakakibara and Hanada, 2011). Acceptor molecules utilised by UGTs include flavonoids, phenolic acids, plant hormones such as IAA, xenobiotics, and other small molecules such as alkaloids, terpenoids and cytokinins (Bowles et al., 2005). While UGTs often transfer glycosyl residues to various endogenous compounds, a variety of exogenous xenobiotics, including herbicides, can also act as the acceptor molecules of UGTs. The addition of a glycosyl residue to a herbicide has the general effect of making it more hydrophilic and less toxic. Therefore, mutations in genes coding for UGTs that enable them to glycosylate herbicides, or mutations that cause overexpression of UGTs that can already glycosylate herbicides, can contribute to NTSR in weeds.

UGTs are classified by the presence of a C-terminal consensus sequence, the UDP-glycosyltransferase signature, known as the 'PSPG' motif (Plant Secondary Product Glycosyltransferase). The PSPG motif is made up of 44-amino acids and is thought to be the site of binding of the UDP moiety of the sugar molecule to the enzymes (Caputi et al., 2012; Gloster, 2014; Ross et al., 2001). On the other hand, the N-terminal regions of UGTs is the domain which is involved in the recognition and binding of a diverse array of acceptor molecules. Therefore, the N-terminal domain is generally more variable than the C-terminus (Meech and Mackenzie, 1997). UGTs catalyse O-, S- and N-linkages and can also form different linkages with the same acceptor molecule (Tiwari et al., 2016). For example, *A. thaliana* UGT72B1 has been reported to form an N-glucosidic bond with 3, 4-dichloroaniline and an O-glucosidic bond with 3,4-dihydroxybenzoic acid (Loutre et al., 2003). Previous phylogenetic analyses of plant UGTs has found this protein family to diverge into at least 20 distinct monophyletic groups (Caputi et al., 2012; Ross et al., 2001). However, the taxon sampling of these studies has been limited to angiosperms, and so conclusions cannot be made regarding the evolutionary history of this protein family in the Archaeplastida lineage.

Here I determine the phylogenetic relationships, of both the ATP-binding cassette transporter and uridine diphosphate glycosyltransferase protein superfamilies within the Archaeplastida lineage. I show that the current clade diversity of the ABC transporter superfamily existed in the last common ancestor of the Archaeplastida. This indicates that the ABC transporters diversified into the extant sub-families before the evolution of the last

common ancestor of the Archaeplastida, extant 1.6 billion years ago (Butterfield, 1990; Hackett et al., 2007; Yoon et al., 2004). By contrast UGT proteins are only present in the streptophyte lineage among the Archaeplastida and diversified in the vascular plant lineages. On the other hand, I show that UGTs are present the streptophyte lineage of the Archaeplastida but absent from non-streptophyte lineages. This suggests either that this protein family, which exists in other eukaryotic lineages, was lost from the glaucophyte, rhodophyte and chlorophyte lineages, or that these proteins originated through horizontal gene transfer into the last common ancestor of the streptophytes, extant approximately 1.1-1.2 billion years ago in the Proterozoic era (Becker, 2013; Douzery et al., 2004; Hedges et al., 2004; Herron et al., 2009; Morris et al., 2018; Nie et al., 2020). Furthermore, I show that the UGT family subsequently diversified throughout streptophyte evolution, mostly in the vascular lineage. This analysis suggests that gene superfamilies which evolved in the Proterozoic era, are now undergoing natural selection caused by herbicide application which is leading to the evolution of herbicide-resistant weeds in agricultural systems.

2.3 Materials and Methods

2.3.1 Data Resources

Amino acid sequences of proteins were obtained from the following sources:

Species	**Group**	**Data resource**	**Access**	**Reference**
Arabidopsis thaliana	Angiosperm (eudicot)	TAIR10	https://www.arabidopsis.org/	(TAIR, n.d.)
Oryza sativa	Angiosperm (monocot)	Rice Genome Annotation Project	http://rice.uga.edu	("Rice Genome Annotation Project," n.d.)
Marchantia polymorpha	Liverwort	MarpolBase	http://marchantia.info/	N/A
Anthoceros agrestis	Hornwort	Genome webpage	https://www.hornworts.uzh.ch/en/download.html	(Li et al., 2020)
Klebsormidium nitens	Streptophyte alga	Genome webpage	http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/	(Hori et al., 2014)
Selaginella moellendorffii	Lycophyte	Phytozome 12	https://phytozome.jgi.doe.gov/pz/portal.html	(Phytozome v12.1, n.d.)
Physcomitrium patens	Moss	Phytozome 12	https://phytozome.jgi.doe.gov/pz/portal.html	(Phytozome v12.1, n.d.)
Chlamydomona reinhardtii	Chlorophyte alga	Phytozome 12	https://phytozome.jgi.doe.gov/pz/portal.html	(Phytozome v12.1, n.d.)
Cyanidioschyzon merolae	Rhodophyte alga	Genome webpage	https://www.genome.jp/kegg-bin/show_organism?org=cme	(Matsuzaki et al., 2004)
Cyanophora paradoxa	Glaucophyte alga	PhycoCosm	https://phycocosm.jgi.doe.gov/Cyapar1/Cyapar1.home.html	(Price et al., 2019)

2.3.2 Sequence collection

Family 1 UGT and ABC transporter protein sequences from *A. thaliana* and *O. sativa* ("Rice Genome Annotation Project," n.d.)("The Arabidopsis Information Resource (TAIR)," n.d.), were used to perform BLASTP searches against the predicted proteomes of *S. moellendorffii, M. polymorpha, A. agrestis, P. patens, K. nitens, C. reinhardtii* and *C. merolae*. A minimum E-value threshold of $1e^{-10}$ was employed. The initial list of sequences for each species collected was then used to 'self-blast' the predicted proteomes of that species in order to retrieve sequences belonging to species-specific Family 1 UGT and ABC transporter sub-families. Using InterProScan 92.0 (Blum et al., 2021), each Family 1 UGT sequence was checked for the UDP-glycosyltransferase domain (SSF53756), while each ABC transporter was checked for the presence of the ABC transporter pfam domain (PF12848). All sequences not containing the searched-for domains were removed before further analysis.

2.3.3 Sequence alignment

Collected amino acid sequences were aligned using the programme MAFFT (Katoh et al., 2019) and its FFT-NS-2 algorithm. Alignments were visualised using Unipro ugene v. 45.1 (Okonechnikov et al., 2012). Sequences which were seen to be lacking important functional residues were removed. Sequences were trimmed manually, with large non-conserved regions removed, leaving only mostly conserved regions without gaps.

2.3.4 Phylogenetic analysis

Final trimmed amino acid sequence alignments were used in a maximum-likelihood phylogenetic analysis using the software PHyML 3.0 (Guindon et al., 2010). Parameters used were an estimated gamma distribution, the LG+G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio test (aLRT). Resulting phylogenies were visualised in FigTree v1.4.4 and annotated in Inkscape v1.0.2.

2.4 Results

2.4.1 960 ABC transporter and 552 UGT sequences were identified in the genomes of the 10 species of Archaeplastida

I first set out to investigate the evolution of ATP-binding cassette (ABC) transporter and uridine diphosphate (UDP) glycosltransferase (UGT) gene families in the Archaeplastida lineage. The Archaeplastida is the lineage which includes glaucophyte algae, rhodophyte algae, chlorophyte algae and the streptophytes. ABC-transporter protein-encoding sequences were identified from 10 selected species (Table 2.1), representing distinct Archaeplastida lineages. UGT sequences were identified in the genomes of 6 of these 10 Archaeplastida species. There were 60 ABCs and 3 UGTs in the genome of the streptophyte alga *Klebsormidium nitens*; 100 ABCs and 50 UGTs in the liverwort *Marchantia polymorpha*; 85 ABCs and 15 UGTs in the moss *Physcomitrium patens*; 166 ABCs and 163 UGTs in the lycophyte *Selaginella moellendorffii*; 124 ABCs and 200 UGTs in the monocot *Oryza sativa*; and 120 ABCs and 121 UGTs in the eudicot *Arabidopsis thaliana*. UGT sequences were not recovered from the genomes of the glaucophyte algae *Cyanophora paradoxa,* the rhodophyte alga *Cyanidioschyzon merolae*, the chlorophyte alga *Chlamydomonas reinhardtii*, or the hornwort *Anthoceros agrestis*. While no UGT encoding genes were identified in these species, 111, 17, 47, and 130 ABC transporter coding sequences were identified in these species respectively. Taken together these data indicate that the ABC transporters were present in the genome of the last common ancestor of the Archaeplastida, while UGT encoding genes have been present in the Archaeplastida lineage since the divergence of the streptophytes from the last common ancestor shared with the chlorophytes.

Species	Classification	Genome (Mb)	Protein-coding genes	ABCs	ABCs (% PCG)	UGTs	UGTs (% PCG)	References
Arabidopsis thaliana	Angiosperm eudicot	135	25,498	120	0.471	121	0.475	("Arabidiposis Genome Initiative," 2000)
Oryza sativa	Angiosperm monocot	321	35,681	124	0.339	200	0.561	(Kawahara et al., 2013)
Selaginella moellendorffii	Lycophyte	212.6	22,285	166	0.745	163	0.731	(Banks et al., 2011)
Anthoceros agrestis	Hornwort	133	24,700	130	0.526	0	0	(Li et al., 2020)
Physcomitrella patens	Moss	480	35,938	85	0.237	15	0.042	(Rensing et al., 2008)
Marchantia polymorpha	Liverwort	225.8	19,138	100	0.523	50	0.261	(Bowman et al., 2017)
Klebsormidium nitens	Streptophyte alga	117.1	16,215	60	0.370	3	0.0185	(Hori et al., 2014)
Chlamydomonas reinhardtii	Chlorophyte alga	120	15,143	47	0.310	0	0	(Merchant et al., 2007)
Cyanidioschyzon merolae	Rhodophyte alga	16.5	5,331	17	0.319	0	0	(Matsuzaki et al., 2004)
Cyanophora paradoxa	Glaucophyte alga	98.4	25,518	111	0.435	0	0	(Price et al., 2019)

Table 2.1. List of members of Archaeplastida used in the analysis
Including their phylum (Classification), genome size (Genome), total number of protein-coding genes (Protein-coding genes), total number of ABC transporter proteins (ABCs), ABC transporter proteins as a percentage of total protein coding genes (ABCs % PCG), total number of UDP-glycosyltransferase proteins (UGTs), UDP-glycosyltransferase proteins as a percentage of total protein coding genes (UGTs % PCG) and the reference for each genome sequence.

2.4.2 Plant ABC transporters are ancient, and all 12 groups existed in the last common ancestor of the Archaeplastida

To investigate the evolutionary history of the ABC transporter protein family, I generated a multiple sequence alignment of 960 ABC transporter protein sequences, which was used to construct a Maximum Likelihood tree (Fig. 2.1). Alignments were trimmed manually to remove large gaps and only a nucleotide-binding domain (NBD) of each sequence was retained.

The topology of the phylogenetic tree shows that plant ABC transporters diverged into 6 previously identified, monophyletic clades. Each of these monophyletic clades has been classified by a single sub-family of ABC-transporters, with each subfamily designated a letter. For example, the monophyletic clade comprising of ABCA proteins is designated clade A. However, ABC transporter sub-families ABCE and ABCF are sister clades, and together form a single monophyletic clade, designated clade E/F. My analysis revealed the previously defined monophyletic clades A, B, C, D, E/F, G (Fig. 2.1). I also identified 6 previously unreported monophyletic clades that comprise sequences from a single species (*C. paradoxa* – Cp) or two species (*C. paradoxa* and *C. merolae* – CpCm) – Cp1, Cp2, Cp3, Cp4, Cp5, and CpCm1 (Fig. 2.1). ABCI proteins are the only previously reported sub-family that do not form a monophyletic clade. ABCI proteins were found to comprise two distantly related monophyletic groups (and therefore constitute a polyphyletic group) (Fig. 2.1). Sequences from the 4 clades B, C, E/F and G are encoded in the genomes of all 10 species of Archaeplastida used in this analysis. This suggests that these four clades were present in the last common ancestor of the Archaeplastida (Fig. 2.2).

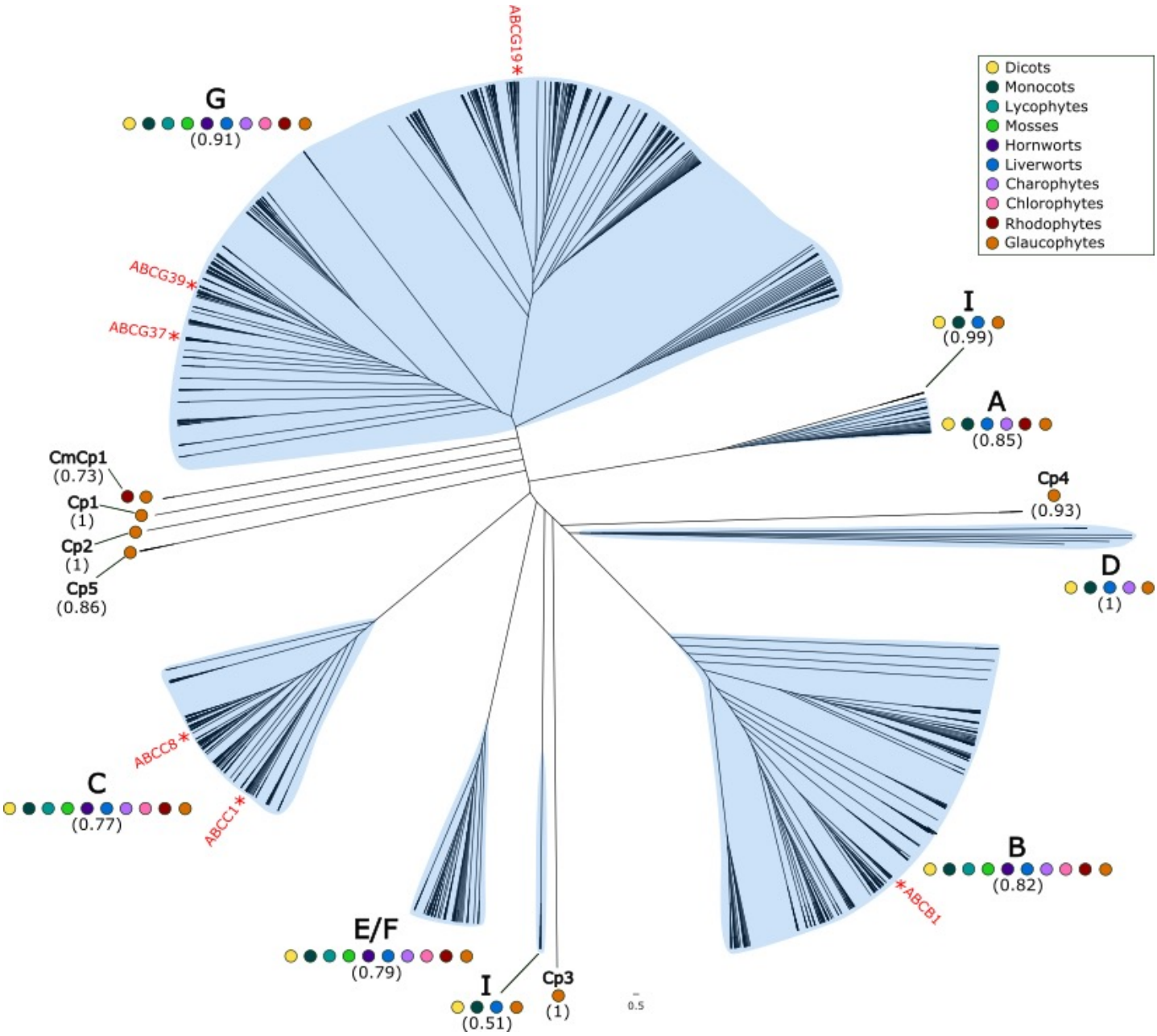


Fig. 2.1. Phylogenetic analysis of ABC transporter protein sequences in the Archaeplastida
Unrooted cladogram of a maximum likelihood (ML) analysis of Archaeplastida ABC transporter proteins conducted by PhyML 3.0 (Guindon et al., 2010) using an estimated gamma distribution parameter, the LG_G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio (aLRT) test. Support values from the aLRT test for each clade are presented in parentheses. Protein sequences were aligned using MAFFT with the L-INS-I algorithm. ABC transporter classes are indicated by blue highlighting and letters. Coloured dots indicate the presence of sequences from different species in each clade. *A. thaliana* (yellow); *O. sativa* (dark green); *S. moellendorffii* (light blue); *P. patens* (light green); *A. agrestis* (dark purple); *M. polymorpha* (blue); *K. nitens* (light purple); *C. reinhardtii* (pink); *C. merolae* (maroon); *C. paradoxa* (orange). Confirmed NTSR genes or their *A. thaliana* homolog reported in Table 2.3 are indicated.

The presence of an ABCA protein sequences in the rhodophyte *C. merolae* and the glaucophyte *C. paradoxa* suggests clade A is ancient and was present in the last common ancestor of the Archaeplastida. However, ABCA proteins are missing from the genomes of 4 species – *C. reinhardtii, A. agrestis, P. patens* and *S. moellendorffii*. The absence of ABCA proteins from these lineages likely represent independent gene loss events, following the divergence of these lineages from other land plants. There are no ABCA proteins encoded in the genome of the chlorophyte alga *C. reinhardtii*, but ABCA proteins are encoded by the streptophyte alga *K. nitens*. This suggests a loss of these proteins in the chlorophyte algae lineage sometime after its divergence from the last common ancestor of the chlorophytes and streptophytes. ABCA proteins are present in the genome of the liverwort *M. polymorpha* but absent from the other two bryophyte species included in the analysis – the hornwort *A. agrestis* and the moss *P. patens*. If hornworts are sister to the mosses and liverworts (Li et al., 2020; J. Zhang et al., 2020), this topology suggests ABCA proteins were lost from the hornwort lineage following its divergence from the moss and liverwort lineages. ABCA proteins were then likely subsequently lost in the moss lineage following its divergence from the liverworts. ABCA proteins were absent from the genome of *S. moellendorffii* but present in the other 2 vascular plants included in the analysis – *A. thaliana* and *O. sativa*. This indicates a loss of ABCA proteins in the lycophyte lineage following its divergence from the angiosperm lineage (Fig. 2.2). The absence of ABCA proteins in some lineages within the Archaeplastida is likely due to ABCA proteins being secondarily lost in these lineages, and hence ABCA proteins were likely already present in the last common ancestor of the Archaeplastida.

ABCD proteins were present in 5 lineages of the Archaeplastida, the dicots, monocots, liverworts, streptophyte algae and glaucophyte algae. The presence of ABCD proteins in the genome of the glaucophyte *C. paradoxa* suggest that ABCD proteins had already diverged from other ABC transporters in the last common ancestor of the Archaeplastida. However, ABCD sequences were not found in *C. merolae* or *C. reinhardtii*, suggesting loss of the ABCD subfamily in the rhodophyte and chlorophyte algal lineages. ABCD proteins were missing from the genomes of *A. agrestis, P. patens* and *S. moellendorffii,* as observed for ABCA proteins. Their absence from these lineages suggests recent gene losses in the hornworts, mosses, and lycophytes respectively. Angiosperm genomes were found to encode ABCD

proteins, however, the other vascular plant included in the analysis, *S. moellendorffii,* does not encode any ABCD proteins, suggesting that this clade was lost in this lineage following the divergence of the lycophyte lineage from the other vascular plants (Fig. 2.2). ABCD proteins were therefore likely present in the last common ancestor of the Archaeplastida lineage, but were subsequently lost in the rhodophyte, chlorophyte, hornwort, moss and lycophyte lineages.

Many putative ABCI encoding genes were identified in the genomes of the plant species in this analysis. However, uniquely amongst ABC transporter proteins, each subunit of an ABCI protein is encoded by a separate gene. Therefore, many ABCI protein sequences identified were removed from the alignments generated as they did not align with the NBD of other ABC-transporters. In the phylogenetic analysis of the retained sequences, 2 monophyletic ABCI clades were defined. One was sister to the monophyletic ABCA protein clade, while the other is sister to a monophyletic clade that includes ABCD and ABCB proteins. Each of the ABCI clades comprise sequences from *A. thaliana*, *O. sativa, M. polymorpha* and *C. paradoxa.* ABCI proteins-encoding genes in the genome of *C. paradoxa* and the 3 previously mentioned land plant species is evidence that they were present in the last common ancestor of the Archaeplastida. However, absence of ABCI protein-encoding genes in the genomes of *C. merolae, C. reinhardtii* and *K. nitens* suggests that ABCI protein-encoding genes have been lost in these algal lineages. Furthermore, the tree topology shows that losses of ABCI proteins have occurred in the land plants. The absence of ABCI proteins-encoding genes in *A. agrestis, P. patens* and *S. moellendorffii* suggest a loss of ABCI proteins in the hornwort, moss and lycophyte lineages (Fig. 2.2). However, since many putative ABCI proteins were removed from the analysis, I cannot be entirely certain that members of this subfamily have been entirely lost from these lineages mentioned. The evolution of the ABCI gene clade may be more difficult to resolve compared to other ABC transporter clades, due to their unique property of individual genes encoding each protein subunit.

Taken together the data show that ABC transporters are conserved among the Archaeplastida. There are sequences from the 10 sampled species from the Archaeplastida that constitute four previously identified monophyletic clades, B, C, E/F and G (Fig. 2.1). This suggests that ABC transporters had diverged into these four subfamilies before the divergence of the Archaeplastida lineages and were therefore present in the last common

ancestor of all species included in the analysis. I also identified 6 other ancient monophyletic clades of ABC transporter which constituted sequences that were present in either glaucophyte, or both glaucophyte and rhodophyte algae (Cp1, Cp2, Cp3, Cp4, Cp5 and CpCm1). ABCA, ABCD and ABCI proteins are not present in the genomes of every species of the Archaeplastida analysed here. My phylogenetic analysis revealed that although ABCA proteins were present in the last common ancestor of the Archaeplastida, members of this clade were subsequently lost in the chlorophyte, hornwort, moss and lycophyte lineages. I show that ABCD proteins also had an ancient origin, likely already having diverged in the last common ancestor of the Archaeplastida. ABCD proteins were subsequently lost in the rhodophyte, chlorophyte, hornwort, moss and lycophyte lineages. The ABCI proteins included in this analysis comprise 2, distantly related clades. Conservation of ABCI proteins included in this analysis was limited, however, their presence in *C. paradoxa* suggests an ancient origin of ABCI proteins. No ABCI protein-encoding sequences were found in the rhodophyte, chlorophyte, streptophyte algae, hornwort, moss or lycophyte lineages, which is consistent with the loss of genes encoding these proteins from these lineages. This suggests that ABCI proteins are ancient but have been lost from many lineages during Archaeplastida evolution.

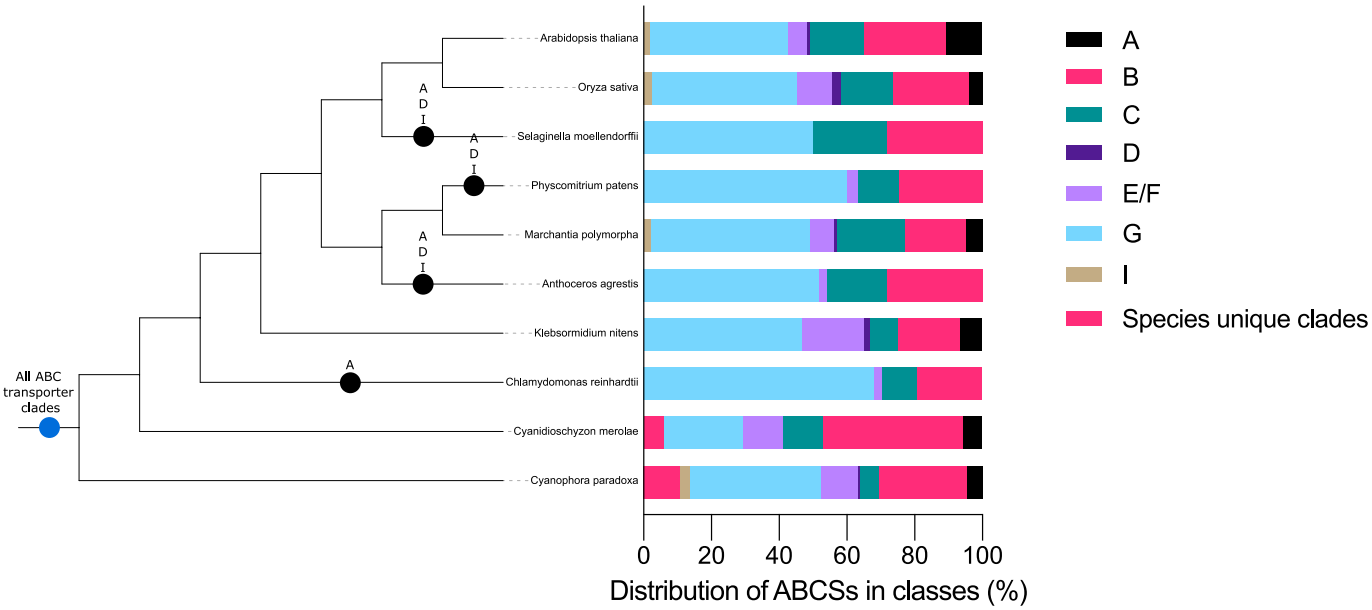


Fig. 2.2. ABC transporters clades were present in the last common ancestor of the Archaeplastida

Cladogram of Archaeplastida showing ABC transporter clade origins and losses during plant evolution. Blue circles represent first appearances of a clade, black circles represent the loss of a clade in a particular lineage. Also shown are the relative distributions of ABC transporters into their respective clades for each species included in this analysis, where each distinct coloured block represents a separate ABC transporter protein clade.

2.4.3 Plant UGTs diversified recently, 14/20 identified groups are only present in Angiosperms

There are currently 45 known families of glycosyltransferases (GTs) in plants (Hansen et al., 2012). Uridine diphosphate glycosyltransferases (UGTs) are the largest glycosyltransferase family and are the only glycosyltransferase family to have been reported to contribute to NTSR. They comprise a monophyletic clade, known as Family 1 GTs. Family 28 GTs are a distinct glycosyltransferase protein family that are related to UGTs (Drula et al., 2022; Fan et al., 2017; Yonekura-Sakakibara and Hanada, 2011; P. Zhang et al., 2020). Family 28 GTs have not been reported to contribute to NTSR. Family 28 GTs were included as an outgroup for this phylogenetic analysis. To investigate the evolutionary history of UGTs in Archaeplastida, phylogenetic trees were constructed using maximum likelihood statistics (Fig. 2.3). This analysis showed that UGTs are only present in the streptophyte lineage of the Archaeplastida. UGTs are present in other eukaryotes, suggesting they were either lost in the glaucophyte, rhodophyte and chlorophyte lineages, or were introduced into the last common ancestor of the streptophyte lineage by horizontal gene transfer. Evidence of horizontal gene transfer of a glycosyltransferase gene has been previously reported from a Verrucomicrobium into the green alga *Picochlorum* SENEW3 (Foflonker et al., 2015). My phylogenetic analysis also shows UGT sequences comprise 20 major monophyletic groups, 16 of which were defined in previous analyses (Caputi et al., 2012; Ross et al., 2001).

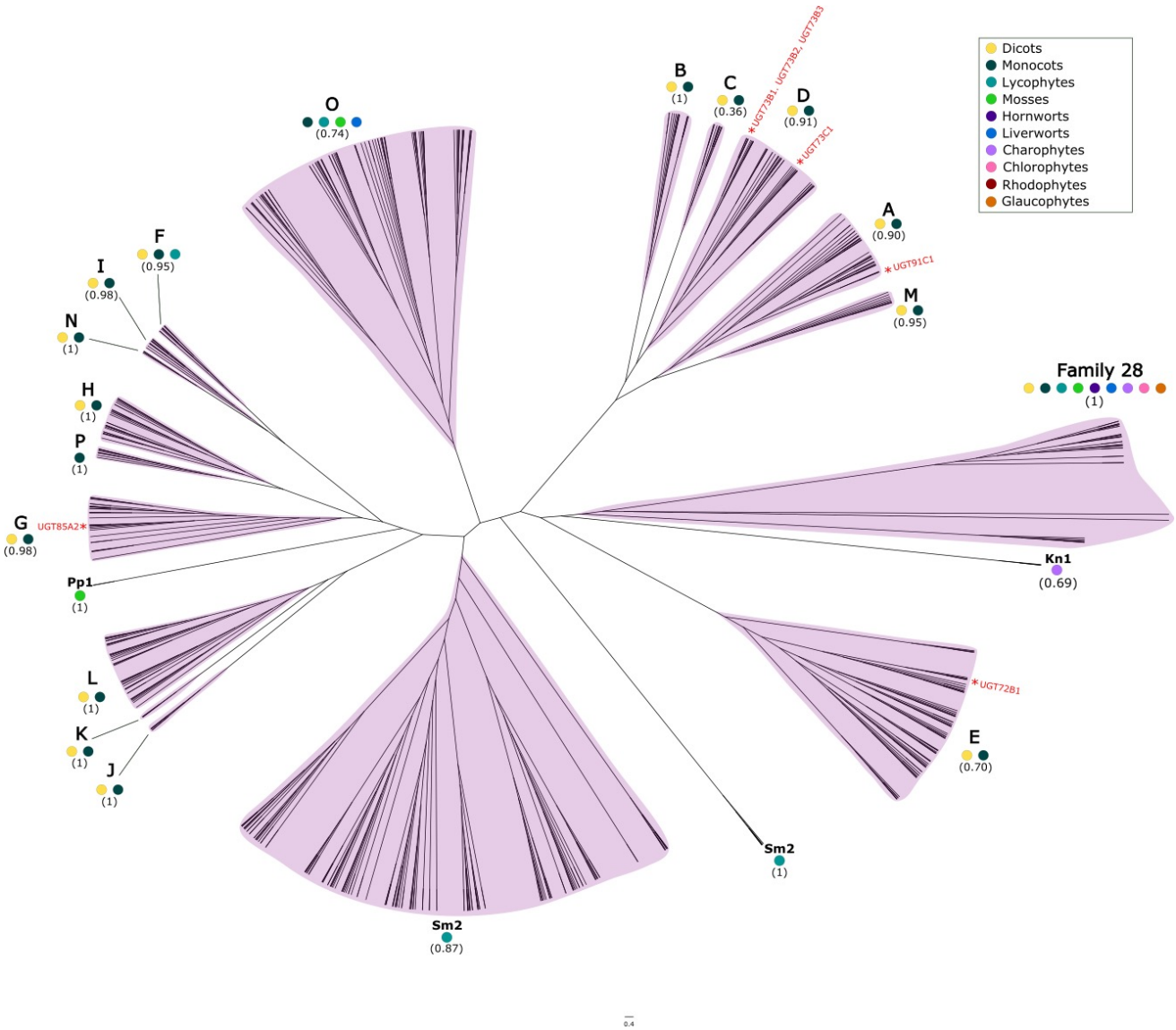


Fig. 2.3. Phylogenetic analysis of UGT protein sequences in the Archaeplastida
Unrooted cladogram of a maximum likelihood (ML) analysis of Archaeplastida UGT proteins conducted by PhyML 3.0 (Guindon et al., 2010) using an estimated gamma distribution parameter, the LG_G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio (aLRT) test. Support values from the aLRT test for each clade are presented in parentheses. Protein sequences were aligned using MAFFT with the L-INS-I algorithm. UGT classes are indicated by pink highlighting and letters. Coloured dots indicate the presence of sequences from different species in each group. *A. thaliana* (yellow); *O. sativa* (dark green); *S. moellendorffii* (light blue); *P. patens* (light green); *A. agrestis* (dark purple); *M. polymorpha* (blue); *K. nitens* (light purple); *C. reinhardtii* (pink); *C. merolae* (maroon); *C. paradoxa* (orange). Confirmed NTSR genes or their *A. thaliana* homolog reported in Table 2.5 are indicated.

34 Family 28 GT sequences were identified in the genomes of the 10 Archaeplastida species, and these were included in the phylogenetic analysis. Family 28 GT sequences were found in 9 of the 10 Archaeplastida species but are missing from the rhodophyte alga *C. merolae*. This is consistent with Family 28 GTs being present in the last common ancestor of the Archaeplastida but subsequently lost in the rhodophyte algae lineage (Fig. 2.4). This likely represents separate origins of Family 28 GTs and UGTs in the Archaeplastida, the latter of which is only present in the streptophyte lineage.

The UGT protein family has gone through many diversification events since the divergence of Family 28 GTs and UGTs from a last common ancestor. I identified 20 distinct monophyletic groups of UGTs, 16 of which had been previously identified. Previously reported groups are named alphabetically. Each of the 4 new groups identified in this analysis are comprised of UGT sequences from a single species and are named according to that species. The streptophyte algae *K. nitens* is the earliest diverging member of the Archaeplastida found with UGT protein-encoding genes in its genome. Three UGTs were identified in *K. nitens* and these formed a single monophyletic group (Kn1), sister to all other UGTs identified in the analysis. This is consistent with UGTs being present in the last common ancestor of the streptophytes since *K. nitens* is the earliest diverging member of the streptophyte lineage included in this analysis. The group O clade comprises 83 UGTs from 4 species of the Archaeplastida – *O. sativa, S. moellendorffii, P. patens* and *M. polymorpha*. Conservation of these sequences in these 4 land plant species and their absence from the genomes of the streptophyte algae indicates that this clade originated after the divergence of land plants from the streptophyte algae. Group O UGTs sequences are not present in the genomes of *A. thaliana* and *A. agrestis*, suggesting a loss of genes in the eudicot lineage after the divergence of *O. sativa* and *A. thaliana* from their last common ancestor, and in the hornwort lineage respectively. UGTs sequences from *O. sativa, S. moellendorffii, P. patens, M. polymorpha* and also *A. thaliana* constitute a large monophyletic clade consisting of UGT protein from groups F, G, H, I, J, K, L, N, P and Pp1. This suggests that this clade originated at approximately the same stage as group O UGTs, and later diverged into each of the 10 groups (Caputi et al., 2012; Ross et al., 2001). This evidence suggests few divergence events of UGT groups occurred during early streptophyte evolution and then more prolific diversification in more recently diverging lineages.

13 of the 20 UGT groups – A, B, C, D, E, G, H, I, J, K, L, M, N – are only present in the genomes of *A. thaliana* and *O. sativa*. This is consistent with these groups being present in the last common ancestor of the angiosperms (Fig. 2.4). Since these sequences are not present in the genome of *S. moellendorffii*, it suggests that they originated after the divergence of the seed plant lineage from the lycophytes. There are group F UGT sequences in each of the 3 vascular plant species included in the analysis, suggesting this group originated in the last common ancestor of the vascular plant lineage. Group F is sister to groups N and I, which comprise only angiosperm sequences. This suggests groups N and I diverged from a common ancestor which they shared with group F. These data show most UGT diversification in extant taxa occurred after the divergence of the seed plants from the lycophytes.

Five groups were defined by sequences from a singles species in this analysis. Group P comprises UGT sequences from only the monocot angiosperm *O. sativa*. There are sequences from only the lycophyte *S. moellendorffii* in groups Sm1 and Sm2, sequences from only the streptophyte algae *K. nitens* in the group Kn1, and there are sequences only from the moss *P. patens* in group Pp1. The formation of these monospecific clades is consistent with recent origins of these sequences but could also be an artifact of poor taxon sampling in this analysis. Group P is only found in *O. sativa* and not the other angiosperm *A. thaliana*, suggesting that group P originated after the divergence of the monocots and the eudicots or was lost in the dicot lineage. Group Sm1 is the largest clade identified in my analysis and comprises 141 UGT sequences. Sm1 sequences were only found in the *S. moellendorffii* genome. This is consistent with this group originating in the lycophyte lineage, following the divergence of lycophytes from other vascular plants. The large number of sequences in this group suggest a prolific expansion of the Sm1 UGTs in the lycophyte lineage. Sequences from group Sm2 were also identified only in *S. moellendorffii*, indicating this group also originated in the lycophyte lineage. However, with 6 sequences in this group, Sm2 did not go through the same extensive expansion as Sm1. Streptophyte algae sequences from *K. nitens* form a monophyletic clade (Kn1) which is sister to all other UGTs. Kn1 likely represents the earliest diverging UGTs identified in this analysis. Two moss sequences form a monophyletic clade (Pp1). Pp1 is sister to a larger monophyletic clade consisting of groups F, G, H, I, P and N. This suggests Pp1 genes share a common ancestor

with sequences from groups F, G, H, I, P and N which that existed before the divergence of bryophytes from other land plants. However, the absence of Pp1 genes in *M. polymorpha* and *A. agrestis* suggest that sequences from this group were lost in these lineages. The origins and losses of these groups in different Archaeplastida lineages are summarised in Fig. 2.4.

In summary, I show UGT sequences are present in streptophyte algae but not in glaucophytes, rhodophytes or chlorophytes; UGTs are encoded in the *K. nitens* genome and in land plants except the hornworts where I propose they were secondarily lost. This indicates that UGTs first arose after the divergence of the streptophyte lineage and continued to diversify following the colonisation of land.

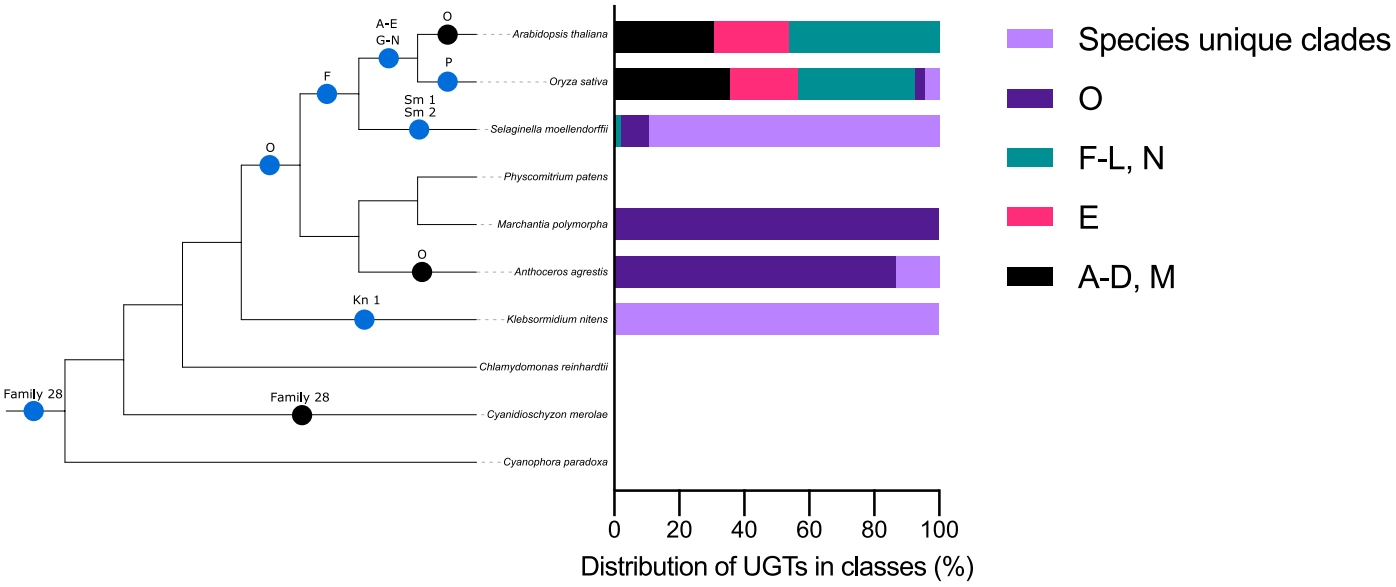


Fig. 2.4. UGTs are present in streptophytes and have diversified into many clades in the vascular lineage Angiosperms
Cladogram of Archaeplastida phylogeny showing UGT clade origins and losses during plant evolution. The origins and losses of Family 28 glycosyltransferases are also shown. Blue circles represent first appearances of a clade, black circles represent the absence of a clade in a particular lineage. Also shown are the relative distributions of UGTs into their respective clades for each species included in this analysis, where each distinct coloured block represents either a separate UGT clade or a group of clades.

2.4.4 Six ABC-transporter sub-families have been associated with non-target site resistance to herbicides

ABC transporters have been implicated in non-target site herbicide resistance (NTSR) of crops and weeds. To identify the clades ABC transporters that are associated with NTSR belong, a literature survey was conducted. This survey identified genes that have been associated with NTSR because the mRNA levels of the ABC transporter mRNA are higher in resistant than sensitive populations. These genes are classified as "candidate NTSR genes" and are listed in (Table 2.2). ABC transporters demonstrated to increase herbicide resistance when expressed at high levels as a transgene or by other experimental methods are classified as "confirmed NTSR genes" and are listed in (Table 2.3).

A total of 41 ABC transporter candidate NTSR genes were identified in a variety of weed and non-weed species from the literature (Table 2.2). They include the grass weeds, Italian ryegrass (*Lolium multiflorum*) (Cechin et al., 2020), red sprangletop (*Leptochloa chinensis*) (Chen et al., 2021), American slough grass (*Beckmannia syzigachne*) (Bai et al., 2020), shortawn foxtail (*Alopecurus aequalis*) (Zhao et al., 2017) and jungle rice (*Echinochloa colona*) (Rangani et al., 2022), the broad leaf (eudicot) weeds flixweed (*Descurainia Sophia*) (Yang et al., 2016), hairy fleabane (*Conyza bonariensis*) (Piasecki et al., 2019), giant chickweed (*Myosoton aquaticum*) (Liu et al., 2018) (Bai et al., 2018), and Palmer amaranth (*Amarnthus palmeri*) (Salas-Perez et al., 2018) and the crop rice (*Oryza sativa*) (Zhao et al., 2023). The candidate NTSR genes were identified in populations resistance to 5 different HRAC mode of action herbicide classes, including acetolactate synthase inhibitors, acetyl coA carboxylase inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors, glutamine synthetase inhibitors and auxin mimic/cellulose synthesis inhibitors.

There are 12 monophyletic clades of ABC transporters (A, B, C, D, E/F, G, Cp1, Cp2, Cp3, Cp4, Cp5 and CpCm1). Candidate NTSR genes encoding ABC transporters are found in sub-families ABCA, ABCB, ABCC, ABCF, ABCG and ABCI (Table 2.2). Thirteen ABCB and 18 ABCC members were identified as candidate NTSR genes. ABCG proteins constitute by far the largest monophyletic clade I identified in the phylogenetic analysis in Section 2.4.2, representing approximately 2/3 of all the ABC transporters included. However, relatively few

ABCG transporters were identified as candidate NTSR genes. From 4 species, 6 distinct ABCG candidate NTSR genes were identified – ABCG7, ABCG8, ABCG11, ABCG14, ABCG39 and ABCG48. A single candidate NTSR gene was found in the literature from each of the groups ABCA, ABCF and ABCI. Taken together, these data suggest candidate NTSR genes encoding ABC transporters belong to clades that were present in the last common ancestor of the Archaeplastida and that the frequency of candidate NTSR genes is not proportional to ABC transporter clade size.

Six ABC transporter proteins have been experimentally confirmed to confer NTSR in plants (Table 2.3). These 6 genes belong to 3 different clades – ABCB, ABCC and ABCG. These clades comprise sequences from all 10 species of Archaeplastida included in the analysis, and therefore had already diverged in the last common ancestor of the Archaeplastida. It can be concluded that ABC transporters that have been confirmed to contribute to NTSR belong to clades which were already present in the common ancestor of the Archaeplastida. The ABC transporters confirmed to confer resistance to herbicides did so to a variety of HRAC mode of action groups, including very long-chain fatty acid inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors, auxin mimics, microtubule assembly inhibitors, PSI inhibitors, and a herbicide with an unknown mode of action (MSMA). I therefore show that ABC transporter proteins within the same monophyletic clade confer resistance to herbicides with different modes of action.

SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE	HRAC CLASS	REFERENCE
ABCA	LmABCA7 (Italian ryegrass)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Cechin et al., 2020)
ABCB	CbABCB1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	MaABCB2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
	MaABCB2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	LcABCB2 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	(Chen et al., 2021)
	ApABCB2 (Palmer amaranth)	Up-regulated in NTSR plants compared to sensitive plants	glufosinate	10	(Salas-Perez et al., 2018)
	OsABCB4 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	BsABCB10 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	(Bai et al., 2020)
	AaABCB11 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	OsABCB11 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	MaABCB19 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	CbABCB25 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	CbABCB27 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	MaABCB29 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
ABCC	DsABCC1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
	CbABCC2 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	MaABCC3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
	MaABCC3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	CbABCC4 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	AaABCC8 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	MaABCC8 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
	CbABCC8 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	LcABCB8 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	(Chen et al., 2021)
	CbABCC9 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	CbABCC10 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	MaABCC10 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
	MaABCC10 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	OsABCB10 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	CbABCC12 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	CbABCC13 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	CbABCC14 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	LcABCB15 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	(Chen et al., 2021)
ABCF	DsABCF5 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
ABCG	CbABCG7 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	MaABCG8 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	OsABCG11 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	MaABCG14 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Liu et al., 2018)
	MaABCG14 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	(Bai et al., 2018)
	OsABCG39 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	EcABCG48 (jungle rice)	regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	(Rangani et al., 2022)
ABCI	CbABCI17 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)

Table 2.2. Plant ABC transporter proteins that have been associated with resistance to herbicides are found within 6 classes identified in this analysis

SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE/ XENOBIOTIC	HRAC CLASS	REFERENCE
ABCB	AtABCB1 (*A. thaliana*)	Endogenous overexpression gives resistance to multiple herbiciides	dicamba, oryzalin, pendimethalin, MSMA	4, 3, 3, 0	(Windsor et al., 2003)
ABCC	AtABCC1 (*A. thaliana*)	Show this gene encodes a protein capable of transporting GS-conjugated metolachlor	metolachlor	15	(Lu et al., 1997)
ABCC	EcABCC8 (jungle rice)	Causes resistance when overexpressed in transgenic rice	glyphosate	9	(Pan et al., 2021)
ABCG	AtABCG19 (*A. thaliana*)	Endogenous overexpression gives resistance	kanamycin	N/A	(Mentewab et al., 2005)
ABCG	AtABCG37 (*A. thaliana*)	GOF mutant more resistant, LOF mutant more sensitive	2,4-D	4	(Ito and Gray, 2006)
ABCG	AtABCG39 (*A. thaliana*)	LOF mutant is resistant	paraquat	22	(Xi et al., 2012)

Table 2.3. Plant ABC transporter proteins that have been empirically shown to confer resistance or tolerance to xenobiotics are found in 3 classes identified in this analysis
The position of these genes or their *A. thaliana* homologs has been indicated in the phylogeny in Fig. 2.1.

2.4.5 UGT groups associated with herbicide resistance evolved after the divergence of Angiosperms from the Archaeplastida lineage

Many UGT genes have also been reported to contribute to NTSR in populations of resistant weeds in the field. I identified a total of 34 candidate NTSR UGT genes in a variety of weed and non-weed species from the literature (Table 2.4). They include the grass weeds, Italian ryegrass (*Lolium multiflorum*) (Cechin et al., 2020), American slough grass (*Beckmannia syzigachne*) (Bai et al., 2020) (Pan et al., 2016), *Brachypodium hybridum* (Matzrafi et al., 2017), shortawn foxtail (*Alopecurus aequalis*) (Zhao et al., 2017), Tausch's goat grass (*Aegilops tauschii*) (Zhang et al., 2022), and jungle rice (*Echinochloa colona*) (Rangani et al., 2022), the broad leaf (eudicot) weeds flixweed (*Descurainia Sophia*) (Yang et al., 2016), hairy fleabane (*Conyza bonariensis*) (Piasecki et al., 2019), and giant chickweed (*Myosoton aquaticum*) (Liu et al., 2018) (Bai et al., 2018) and the crop rice (*Oryza sativa*) (Zhao et al., 2023). The candidate NTSR genes were identified in populations resistant to 4 classes of HRAC mode of action herbicides, including acetolactate synthase inhibitors, acetyl coA carboxylase inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors and auxin mimic/cellulose synthesis inhibitors.

There are 16 monophyletic groups of UGT sequences in angiosperms. Candidate NTSR genes were identified from 11 of these monophyletic groups in angiosperm weeds and crops (Table 2.4). There are 8 candidate NTSR group D UGT genes in 7 different species. These represent 7 distinct genes, UGT73B3, UGT73B5, UGT73C, UGT73C1, UGT73C5, UGT73D1 and UGT73E1. UGT73B5 was identified in both *B. hybridum* and *A. aequalis.* A single candidate NTSR gene was identified in flixweed, UGT71B1, a group E UGT. Group E is the second largest UGT group. No candidate NTSR UGT sequences were identified from group O, which is the largest UGT group in which angiosperm sequences were identified. These data indicate UGT proteins that detoxify herbicides belong to clades that were present in the last common ancestor of the Angiosperms.

Seven UGT genes, from 4 groups – A, D, E and G – have been experimentally confirmed to confer NTSR in plants (Table 2.5). There are genes from only two species, the angiosperms *A. thaliana* and *O. sativa,* in each of these groups in my phylogenetic analysis in Section 2.4.3.

This suggests that these groups were present in the last common ancestor and evolved following the divergence of the angiosperms from the lycophytes. 6 of the 7 genes identified confer NTSR to 3 HRAC mode of action classes, HPPD inhibitors, acetyl coA inhibitors and PSI inhibitors.

GROUP	FAMILY + SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE	HRAC CLASS	REFERENCE
A	79	LmUGT79 (Italian ryegrass)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Cechin et al., 2020)
	79	OsUGT79 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
	79B	DsUGT79B6 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
	91A	DsUGT91A1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
B	89A	CbUGT89A (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
C	90A	BsUGT90A1 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	(Bai et al., 2020)
	71C	DsUGT71C2 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
D	73B	MaUGT73B3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Liu et al., 2018)
	73B	BhUGT73B5 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	(Matzrafi et al., 2017)
	73B	AaUGT73B5 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	73C	AtUGT73C (Tausch's goat grass)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhang et al., 2022)
	73C	BsUGT73C1 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	(Pan et al., 2016)
	73C	AaUGT73C5 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	73D	EcUGT73D1 (jungle rice)	Up-regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	(Rangani et al., 2022)
	73E	CbUGT73E1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
E	71B	DsUGT71B1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
	88A	BhUGT88A1 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	(Matzrafi et al., 2017)
F	78D	MaUGT78D2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Liu et al., 2018)
	78D	MaUGT78D2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Bai et al., 2018)
	78D	DsUGT78D4 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
G	85A	BhUGT85A2 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	(Matzrafi et al., 2017)
	85A	BsUGT85A2 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	(Pan et al., 2016)
	85C	CbUGT85C1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
	85C	CbUGT85C2 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
H	76B	CbUGT76B1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	(Piasecki et al., 2019)
I	83A	AaUGT83A1 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	83A	BsUGT83A1 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	(Bai et al., 2020)
	83A	OsUGT83A1 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	fenoxaprop-P-ethyl	1	(Zhao et al., 2023)
L	74D	DsUGT74D1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
	75D	AaUGT75D1 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)
	75D	EcUGT75D1 (jungle rice)	Up-regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	(Rangani et al., 2022)
	84A	DsUGT84A1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	(Yang et al., 2016)
M	92A	AaUGT92A1 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	(Zhao et al., 2017)

Table 2.4. Plant UGTs associated with herbicide resistance are found in 11 classes I have identified in this analysis

GROUP	FAMILY + SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE/ XENOBIOTIC	HRAC CLASS	REFERENCE
A	91C	AtUGT91C1 (*A. thaliana*)	Demonstrated glucosylation of sulcotrione and herbicide resistance in overexpression lines	sulcotrione	27	(Huang et al., 2021)
E	72B	AtUGT72B1 (*A. thaliana*)	Knockout plants no longer conjugated these chemicals	3,4-dichloroaniline (DCA) and 2,4,5-trichlorophenol (TCP)	N/A	(Brazier-Hicks and Edwards, 2005)
D	73C	BsUGT73C1 (American slough grass)	Identified in RNA-seq experiment – causes resistance when expressed in *B. distachyon*	fenoxaprop-p-ethyl	1	(Pan et al., 2016)
	73B	AtUGT73B1 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	(Lira et al., 2006)
	73B	AtUGT73B2 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	(Lira et al., 2006)
	73B	AtUGT73B3 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	(Lira et al., 2006)
G	85A	BsUGT85A2 (American slough grass)	Identified in RNA-seq experiment – causes resistance when expressed in *B. distachyon*	fenoxaprop-p-ethyl	1	(Pan et al., 2016)

Table 2.5. Plant UGTs that have been empirically shown to confer resistance in plants belong to 4 classes identified in this analysis
The position of these genes or their *A. thaliana* homologs has been indicated in the phylogeny in Fig. 2.3.

2.5 Discussion

Herbicide use has increased continuously since the 1940s, leading to the evolution of herbicide resistance in weeds through selection of naturally occurring alleles that confer herbicide resistance. Overexpression of genes encoding some members of the ATP-binding cassette (ABC) transporter family and uridine diphosphate (UDP) glycosyltransferase (UGT) family in weeds confers herbicide non-target site resistance (NTSR) (Délye et al., 2015; Gaines et al., 2020; Powles and Yu, 2010). ABC transporters are ATP-driven pumps that transport molecules across membranes, while UGT enzymes catalyse the metabolism of a diversity of compounds by transferring a sugar moiety to an acceptor molecule. By carrying out phylogenetic analyses, I set out to resolve the evolutionary relationships of these protein families, and to classify those genes associated with NTSR. I compared sequences from 10 species which represent distinct lineages of the Archaeplastida and generated phylogenetic trees. From the topology of these trees, it can be concluded that ABC transporters were present in the last common ancestor of the Archaeplastida. There are 12 monophyletic clades of ABC transporters (A, B, C, D, E/F, G, Cp1, Cp2, Cp3, Cp4, Cp5, and CpCm1) and one polyphyletic clade (I), all of which existed in the last common ancestor of the Archaeplastida. UGT enzymes existed in the last common ancestor of the streptophytes and have since diversified into 20 monophyletic clades (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Kn1, Pp1, Sm1, and Sm2). Overall, the phylogenetic analysis shows that ABC transporters were present in the last common ancestor of the Archaeplastida, while UGTs proteins are only present in the streptophyte lineage among the Archaeplastida.

In plants, ABC transporters transport of a variety of molecules across membranes (Higgins, 1992; Kang et al., 2011). ABC transporters confer herbicide resistance by sequestering herbicides and their metabolites in the vacuole or extracellular spaces, where they cannot interact with their target (Ge et al., 2010). I hypothesised that a high diversity of ABC transporters among the Archaeplastida may facilitate the high diversity of substrates which can be recognised and transported by ABC transporters. However, I have shown that closely related ABC transporters, within the same clade, can be associated with resistance to a diverse range of herbicide mode of actions. For example, proteins from clade A are associated with resistance to 6 different herbicides, from 4 different HRAC mode of action

groups (HRAC group 1, 2, 9 and 10). This suggests that ABC transporter diversity is not necessary for a diversity of molecules which can be recognised and transported. Nonetheless, this diversity may play a role diversity of ABC transporter localisation, and hence in herbicide detoxification by allowing ABC transporters to sequester chemicals in various locations. These locations include the vacuole and extracellular spaces, where herbicides can be sequestered and are no longer toxic to the plant.

ABC transporter clades in extant plants are ancient and I show that the common ancestor of the Archaeplastida, which existed around 1.6 billion years ago in the Proterozoic era, contained all clades of ABC transporters identified in my analysis (Archibald, 2015; Hedges et al., 2004; Keeling, 2013). These ABC transporters evolved during the Proterozoic to carry out various cellular functions, unrelated to herbicide resistance. However, certain ABC transporter genes are now being selected in modern agricultural systems for a new purpose – conferring resistance to herbicides. For example, ABCB1, a P-glycoprotein, which has been shown to cause resistance to dicamba, oryzalin, pendimethalin and MSMA when endogenously overexpressed in *A. thaliana* has been characterised (Windsor et al., 2003). ABCB1 is involved in the transport of auxin, a crucial plant hormone essential for numerous developmental processes. ABCB1 mediates auxin efflux, regulating its distribution and concentration in different plant tissues. ABCB1 mutants have altered auxin distribution and as a result, show various developmental defects (Blakeslee et al., 2007; Geisler et al., 2005; Geisler and Murphy, 2006). Herbicide resistance evolving through the altered function of genes that evolved to carry out an entirely different function, can be considered exaptation (Gould and Vrba, 1982). Therefore, I argue that the involvement of ABC transporters in herbicide resistance is an example of exaptation.

UGT enzymes catalyse the conjugation of sugars moieties to a variety of acceptor molecules, including herbicides which can lead to detoxification (Dimunová et al., 2022). High sequence diversity among UGTs genes leads to protein structural heterogeneity, which in turn facilitates a wide range of acceptor molecules, including herbicides (Bowles et al., 2005). I have shown that a large proportion of UGT gene diversity evolved in the angiosperm lineage, and this may account for the ability of UGTs to detoxify herbicides. However, I have also shown that UGTs from closely related clades, or even those from within the same clade, may

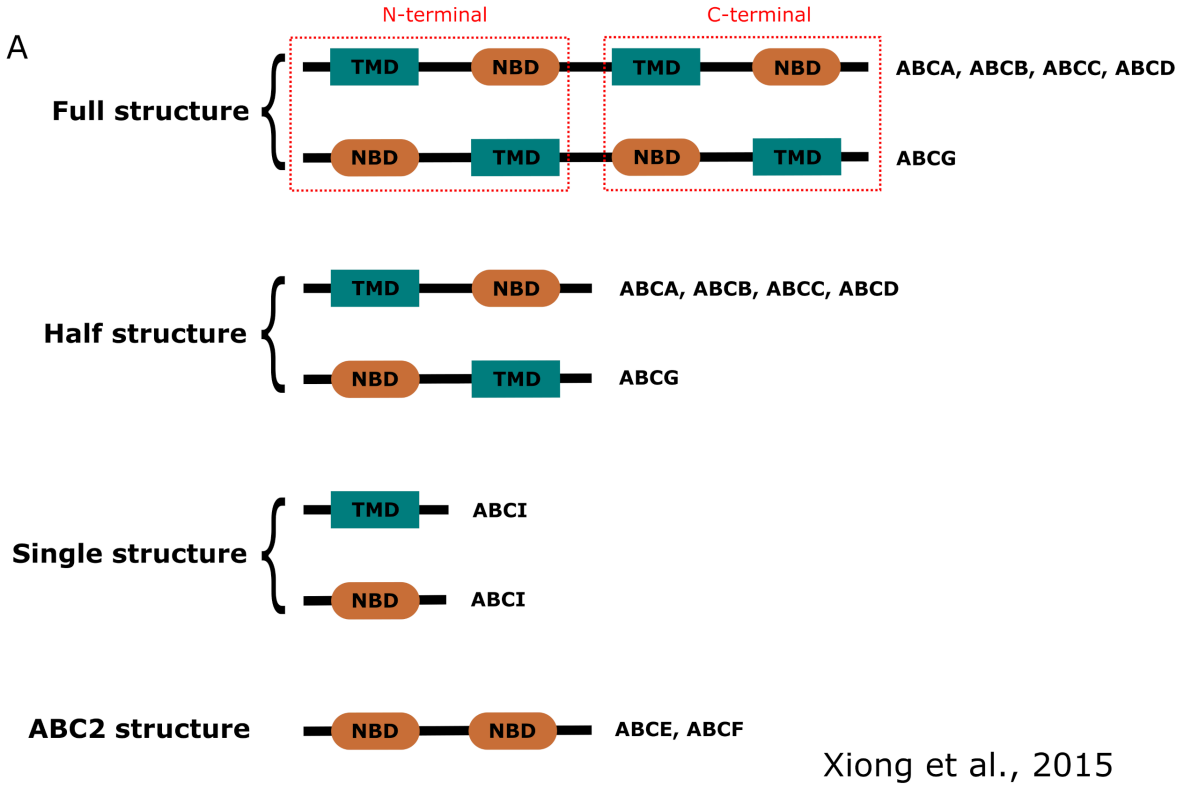
confer resistance to herbicides with a variety of different chemistries and modes of action. For example, UGT proteins from group D have been associated with, or confer resistance to, 7 different herbicides from 5 different HRAC mode of action groups (HRAC group 1, 2, 4/29, 9 and 22). This suggests sufficient variability has also evolved within UGT clades in extant plants to give UGTs the capacity to conjugate sugars to a diverse range of acceptor molecules, which includes herbicides.

UGTs proteins were present in the last common ancestor of the streptophytes, which existed 1.1-1.2 billion years ago, in the Proterozoic era (Becker and Marin, 2009; Delwiche and Cooper, 2015; Leliaert et al., 2012). However, UGTs are not present in the earlier diverging glaucophyte, rhodophyte or chlorophyte lineages. Since UGTs are present in other eukaryotes and some bacteria, their absence in non-streptophyte algal lineages suggests either a loss of UGT genes in these lineages or that UGTs were acquired in the last common ancestor of the streptophyte lineage through horizontal gene transfer. My data suggest that there was little diversification of UGT sequences until after the vascular plants diverged from other streptophytes in the Silurian period of the Paleozoic era (420-450 million years ago) (Becker and Marin, 2009; Delwiche and Cooper, 2015; Kenrick and Crane, 1997; Leliaert et al., 2012). 17 of the 20 UGT clades I identified, evolved in the vascular lineage. 14 of these groups are present only in angiosperms, suggesting they evolved after the divergence of seed plants from lycophytes, which occurred in the Devonian period of the Paleozoic era (360-400 million years ago) (Jiao et al., 2011; Kenrick and Strullu-Derrien, 2014). This diversification of UGTs likely accompanied an increase in morphological complexity of plants and the colonisation of a wider range of terrestrial habitats. UGTs were likely selected to carry out a variety of cellular functions in streptophytes, during their colonisation of land, and subsequent evolution of the vascular plants. Over the last 50 years, they have been selected for herbicide resistance in the agricultural environment.

I have shown that both ABC transporters and UGT enzymes evolved in the Proterozoic era. ABC transporters genes were present in the last common ancestor of the Archaeplastida and had already diversified into many clades in this lineage. Conversely, my data shows that UGTs genes are only present in the streptophyte lineage and diversified mainly in the vascular lineage among plants. The clade diversity of ABC transporters and UGT enzymes in extant

plants may contribute to their ability to confer herbicide resistance. I propose that ABC transporters and UGTs were selected to carry out a variety of cellular functions in the Proterozoic, and now some members of these protein families are being selected to carry out an entirely different function – conferring resistance to herbicides. ABC transporters and UGTs conferring herbicide resistance is therefore an example of exaptation.

2.6 Supplementary Data



Supplementary Fig. S2.1. Protein features of ABC transporters and UDP-glycosyltransferases in plants.
Diagrams of the typical ABC transporter (A) and UDP-glycosyltransferases (B). (A) shows the domain architecture of ABC transporters from each of the major sub-families in plants. The diagram is adapted from Xiong et al., 2015. (B) shows α-helices (green), β-strands (purple) and residues reported to form part of the acceptor pocket (red dashed lines). The PSPG motif reported to interact with the UDP-sugar donor is highlighted in grey. The structure is basted on that of AtUGT72B1 reported in Osmani et al., 2009.

ABC clade	A	B	C	D	E/F	G	I	CmCp1	Cp1	Cp2	Cp3	Cp4	Cp5	Total
Arabidopsis thaliana	13	29	19	1	7	49	2	0	0	0	0	0	0	**120**
Oryza sativa	5	28	19	3	13	53	3	0	0	0	0	0	0	**124**
Selaginella moellendorffi	0	47	36	0	0	83	0	0	0	0	0	0	0	**166**
Anthoceros agrestis	0	32	16	0	4	78	0	0	0	0	0	0	0	**130**
Marchantia polymorpha	5	18	20	1	7	47	2	0	0	0	0	0	0	**100**
Physcomitrium patens	0	24	15	0	2	44	0	0	0	0	0	0	0	**85**
Klebsormidium nitens	4	11	5	1	11	28	0	0	0	0	0	0	0	**60**
Chlamydomonas reinhardtii	0	9	5	0	1	32	0	0	0	0	0	0	0	**47**
Cyanidioschyzon merolae	1	7	2	0	2	4	0	1	0	0	0	0	0	**17**
Cyanophora paradoxa	5	29	6	1	12	43	3	1	1	1	1	3	5	**111**
Total	**33**	**234**	**143**	**7**	**59**	**461**	**10**	**2**	**1**	**1**	**1**	**3**	**5**	**960**

Supplementary Table S2.1. Number of ABC proteins identified in each clade.

UGT clade	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Sm1	Sm2	Pp1	Kn1	Total UGT	Total Family 28
Arabidopsis thaliana	16	4	3	13	28	4	7	21	1	2	2	18	1	1	0	0	0	0	0	0	**121**	**2**
Oryza sativa	17	10	8	28	42	3	23	7	10	3	1	22	8	3	6	9	0	0	0	0	**200**	**6**
Selaginella moellendorffi	1	0	0	0	0	2	0	0	0	0	0	0	0	0	14	0	141	5	0	0	**163**	**6**
Anthoceros agrestis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**6**
Marchantia polymorpha	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	**50**	**1**
Physcomitrium patens	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	2	0	**15**	**5**
Klebsormidium nitens	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	**3**	**4**
Chlamydomonas reinhardtii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**1**
Cyanidioschyzon merolae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0**
Cyanophora paradoxa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**3**
Total	**34**	**14**	**11**	**41**	**70**	**9**	**30**	**28**	**11**	**5**	**3**	**40**	**9**	**4**	**83**	**9**	**141**	**5**	**2**	**3**	**552**	**34**

Supplementary Table S2.2. Number of UGT proteins identified in each clade. Number of Family 28 glycosyltransferases identified in each species is also given.

Chapter 3: *Marchantia polymorpha* is glyphosate-tolerant but lacks a glyphosate-tolerant form of the EPSPS enzyme

3.1. Abstract

Herbicides are a key component of modern agricultural practices and are needed to maintain crop yields. However, some of the weeds that herbicides target have evolved resistance to these chemicals. Many weeds have evolved resistance to the herbicide glyphosate but few, if any, angiosperm plants are naturally tolerant to glyphosate. Herbicide resistance can evolve through mutations affecting the herbicides target site (target-site resistance, TSR) or by mechanisms that reduce herbicide exposure to the target or minimise its toxic impact (non-target site resistance, NTSR). The target of glyphosate is the shikimate pathway enzyme 5-enolpyruvylhikimate-3-phosphate synthase (EPSPS). Here, I show that the liverwort *Marchantia polymorpha* is more than 12 times as tolerant to glyphosate as the angiosperm *Arabidopsis thaliana*. I show that the sequence and structure of EPSPS is highly conserved among land plants. However, through an *in silico* molecular docking analysis, I predict that glyphosate does not efficiently bind to the EPSPS of *M. polymorpha*, potentially contributing to target-site tolerance. In contrast, using enzyme activity assays, I demonstrate that the *M. polymorpha* EPSPS is equally sensitive to glyphosate as other land plant EPSPSs, which are approximately 500-fold more sensitive to glyphosate than EPSPS enzymes used in glyphosate tolerant crops. Through enzyme kinetic analyses, I then show that the *M. polymorpha* EPSPS enzyme is less catalytically efficient than other land plant EPSPS enzymes. I hypothesise that low catalytic efficiency of *M. polymorpha* could be associated with glyphosate tolerance at the whole plant level. I identify a serine at position 500 in the *M. polymorpha* EPSPS which may contribute to this reduced activity. However, site-directed-mutagenesis of this residue in transgenic *A. thaliana* plants does not support the hypotheses that a serine at this position specifically reduces EPSPS activity or contributes to glyphosate tolerance. I conclude that *M. polymorpha* is tolerant to glyphosate and hypothesise that a novel mechanism may confer glyphosate tolerance in *M. polymorpha*.

3.2. Introduction

Intensive use of herbicides since their widespread adoption into modern agricultural practices has led to the evolution of herbicide resistance in weed populations (Heap, 2023). The evolution of herbicide resistance is driven by selective pressure resulting from herbicide applications, which favours certain mutations in genes, leading to populations of plants becoming more resistant to the herbicide (Délye et al., 2013). Glyphosate is the most used herbicide in the world, and as a result, there are many cases of herbicide resistance against glyphosate (Duke and Powles, 2008; Heap and Duke, 2018). Glyphosate is a non-selective herbicide which targets the 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) – a key enzyme in the shikimate pathway. There are few, if any, examples of plants which are naturally tolerant to glyphosate reported in the literature (Duke, 2018). Some other organisms which encode an EPSPS, such as certain bacteria, are naturally tolerant to glyphosate. For example, the CP4 strain of *Agrobacterium tumefaciens* is resistant to glyphosate, due to encoding a resistant form of the EPSPS enzyme. This enzyme has proven useful by allowing the development of the first glyphosate tolerant crop, soybean, which allowed targeted weed control with glyphosate, without causing crop injury (Padgette et al., 1988). The discovery of a plant with natural glyphosate tolerance may uncover novel resistance mechanisms which could allow for new glyphosate tolerant crop technologies to be developed or help discover potential mechanism of herbicide resistance which may evolve in weed populations in the future.

Herbicide resistance in plants can be classified into two main categories: target-site resistance (TSR) and non-target-site resistance (NTSR). So far, only TSR mechanisms have been introduced into crops by genetic modification in order to make them herbicide tolerant. This is likely due to TSR being monogenic and often causing high levels of resistance, whereas NTSR is often polygenic and the magnitude of resistance from each contributing factor is smaller and additive (Powles and Yu, 2010). However, NTSR does form the basis of the mechanism of selectivity of some herbicides. For example, sulfonylurea herbicides, including metsulfuron-methyl and chlorsulfuron are used to control a wide range of broadleaf and grass weeds among crops and are selective due to the action of cytochrome P450 enzymes in the crop (Dimaano and Iwakami, 2021; Frear, 1995; Usui,

2001). Therefore, it is formally possible that an NTSR mechanism, if conferring a high enough level of resistance, could also be employed to improve selectivity to a broad-spectrum herbicide such as glyphosate.

I have discovered that the liverwort *M. polymorpha* is naturally tolerant to glyphosate. The aim of my project following this discovery is to understand the mechanism by which *M. polymorpha* is naturally tolerant to glyphosate – a largely unique trait among plants. Understanding the mechanism of glyphosate tolerance in *M. polymorpha* represents a rare opportunity to investigate a tolerance mechanism to this herbicide that has not evolved in the field. Therefore, the mechanism by which *M. polymorpha* is tolerant to glyphosate is likely to be novel. This may offer potential avenues for developing new ways of increasing the selectivity of glyphosate in agricultural systems. Furthermore, many of the gene networks of the *M. polymorpha* gametophyte generation are similar to those of dominant sporophytes (Bowman et al., 2017; Ligrone et al., 2012; Pires and Dolan, 2012) and as a result, the mechanism of glyphosate tolerance in *M. polymorpha* may plausibly evolve in weed population in the field. Uncovering this mechanism before it has evolved in weed populations could allow for the prevention of yet another glyphosate resistance case in the field. Using a combination of *in silico, in vitro* and *in vivo* methods, I have collected data which are consistent with *M. polymorpha* not encoding a tolerant form of EPSPS. This suggests that *M. polymorpha* is not tolerant to glyphosate through TSR, but NTSR mechanisms.

3.3. Material and Methods

3.3.1. Phylogenetic analysis of EPSPS genes

EPSPS amino acid sequences were identified through a protein BLASTP search using MpEPSPS (Mp6g04140.1) as a query against reference proteomes of Archaeplastida species, *Saccharomyces cerevisiae*, and *Escherichia coli*. Inclusion criteria were an E-value of less than 1E-10, and confirmation as EPSPS proteins was ensured through a Pfam domain search (El-Gebali et al., 2019).

Confirmed EPSPS protein sequences were aligned in MAFFT using the L-INS-I algorithm (Katoh and Toh, 2010) and visualized in Unipro ugene v. 45.1 (Okonechnikov et al., 2012). Sequences were manually trimmed to retain mostly conserved regions while removing large non-conserved regions and gaps.

The final trimmed amino acid sequence alignment was subjected to maximum-likelihood phylogenetic analysis using PhyML 3.0 with an estimated gamma distribution parameter and the LG model of amino acid substitution (Guindon et al., 2010). Support values were generated using a Chi^2-based approximate likelihood ratio test (aLRT). Resulting phylogenies were visualized in FigTree v1.4.4 (Rambaut, 2007) and Interactive Tree of Life (iTOL) v5 (Letunic and Bork, 2021), followed by annotation in Inkscape v1.0.2.

3.3.2. Homology modelling

Homology models of EPSPS enzyme 3D structures were built using SwissModel and refined in UCSF Chimera (Pettersen et al., 2004), based on quality data from MolProbity and QMEAN servers. Amino acid sequences were obtained from Uniprot or the One Thousand Plant Transcriptomes Initiative database (Carpenter et al., 2019) and submitted to SwissModel. The crystal structure of EPSPS from *Vibrio cholerae* in a closed conformation (Minasov, G. et al., unpublished) was used as a template. After initial modelling, hydrogens and partial charges were added. Models were re-assessed in SwissModel for structural quality, identifying Ramachandran and rotamer outliers. Proline residues with Ramachandran outliers were optimized by adjusting rotamers, while other amino acids underwent selective minimisation (100 steepest descent steps). Rotamer outliers were also corrected, with

refinement continuing until Ramachandran and rotamer outliers were minimised. Final structure minimisation (1000 steepest descent steps) resolved clashes and achieved acceptable MolProbity scores.

3.3.3. Molecular docking

EPSPS homology models were sequentially docked with S3P and glyphosate using BioSolveIT SeeSar software (v.11.0.2, 2021). S3P and glyphosate structures in PDB format were extracted from the EPSPS crystal structure of *Vibrio cholerae* using UCSF Chimera. Models underwent DockPrep with UCSF Chimera before docking. Binding sites were defined based on residues from the *E. coli* crystal structure bound to S3P and glyphosate (Schönbrunn et al., 2001). S3P was docked to each model, and the conformation with the highest affinity was saved as a new PDB file. This file was used for docking glyphosate, selecting the highest-affinity conformation. Protein-ligand complexes were analysed with LigPlot+ (v.2.2.4, 2009) to illustrate interactions and glyphosate's 2D conformation when bound to EPSPS.

3.3.4. Plant lines and growth conditions

Wild-type *M. polymorpha* lines were the laboratory accessions Takaragaike-1 (Tak-1) and Takaragaike-2 (Tak-2) (Ishizaki et al., 2008). The wild-type *A. thaliana* line used was Col-0. *Conocephalum conicum* and *Lunularia cruciata* plants were collected from the wild (Supplementary Fig. S3.5).

Liverwort plants were maintained through asexual propagation of gemmae or thallus excision and grown on solid ½ Gamborg medium with 0.8% agar (Section 3.3.14) in a growth chamber at 23 ℃ under 24-hour 10-30 µmol m^{-2} s^{-1} white light. *A. thaliana* plants were grown with a consistent 16-hour day of 85 µmol m^{-2} s^{-1} white light.

3.3.5 DNA constructs

Wild-type EPSPS genes from *A. thaliana* and *M. polymorpha* were amplified via Phusion high-fidelity polymerase (NEB) using primers designed with BsaI overhangs. After running PCR products on a 1% agarose gel, correctly sized fragments were extracted and purified using the Vienna Biocenter’s in-house gel extraction kit. Fragment sequences were verified

via Sanger sequencing. Purified products were cloned using GreenGate cloning (Lampropoulos et al., 2013) into the entry module pGGI000 and transformed into *E. coli*. Plasmids were extracted and purified using the Vienna Biocenter's in-house kit. These plasmids, along with the promoter (35S) and terminator (RBCS) modules, were combined in a reaction with the destination vector (pGGZ003) and used for site-directed mutagenesis and subsequent experiments. The final vectors were introduced into *A. tumefaciens* through electroporation.

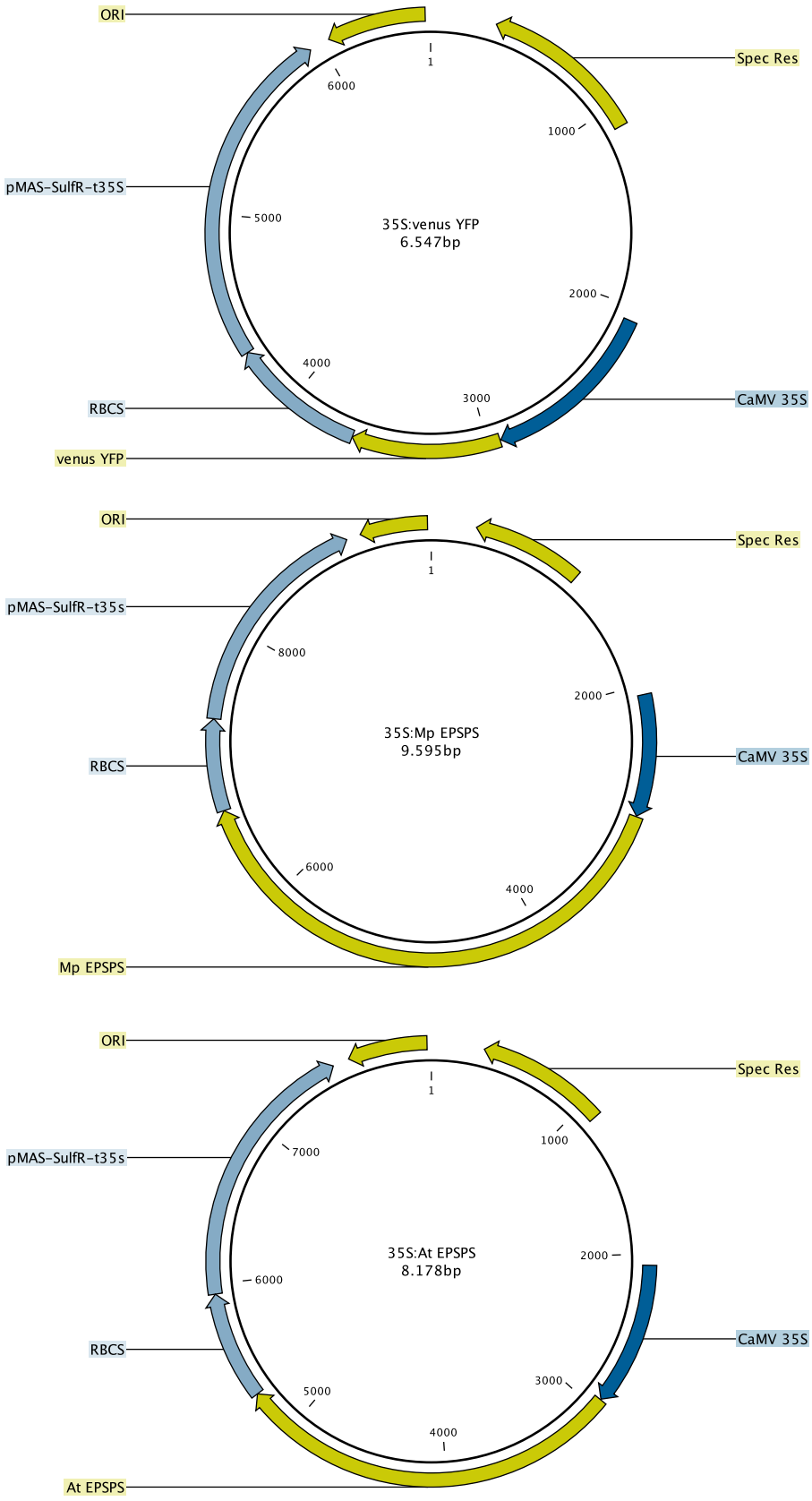


Fig. 3.1. Vector map of *EPSPS* gene expression constructs
Full length DNA sequences of the *EPSPS* genes from *A. thaliana* and *M. polymorpha* were cloned into a vector using GreenGate cloning. These vectors were also used for site-directed mutagenesis.

3.3.6. *Agrobacterium tumefaciens*-mediated transformation of *A. thaliana* plants

To introduce overexpression constructs into *A. thaliana*, *Agrobacterium tumefaciens*-mediated floral dipping was employed. Vectors were transformed into *Agrobacterium* strain GV3101 via electroporation, a strain resistant to rifampicin and gentamycin, with successful transformants showing resistance to spectromycin. Transformed *A. tumefaciens* was plated on solid LB media supplemented with 50 µg ml^{-1} of spectromycin, gentamycin, and rifampicin, and grown for 2 days at 28 ℃ to select colonies that had taken up the overexpression vectors. Single colonies were subsequently grown in 5 ml of liquid LB media supplemented with spectromycin, gentamycin, and rifampicin, and incubated for 2 days with shaking at 150 rpm at 28 ℃. The culture was then used to inoculate 100 ml of liquid LB with the same antibiotics and incubated for a further 2 days. The 100 ml culture was centrifuged, and resulting pellet washed with 20 ml of a 5% sucrose water solution. The pellet was resuspended in 50 ml of 5% sucrose water solution, with the addition of 10 µl of SILWET L-77. Healthy *A. thaliana* Col-0 plants had their flowers submerged in the *Agrobacterium* solution for 5 minutes in a beaker. Plants were then placed in the dark overnight. Seeds from these plants were plated on solid MS media (Section 3.3.14.) supplemented with 7.5 µg ml^{-1} of sulfadiazine. The seeds were stratified for 2 days at 4 ℃ and in the dark before germinating in a chamber with 16 hours of 85 µmol m^{-2} s^{-1} white light and 8 hours of darkness, growing in these conditions for 2 weeks. Selected growing plants were transferred to soil, and homozygous lines for the overexpression constructs were generated.

3.3.7. Herbicide and stock solution preparation

Glyphosate (R1600) was obtained from Honeywell Fluka. Stock solutions were prepared by dissolving glyphosate in MonoQ H_2O.

3.3.8. Dose-response assays

M. polymorpha gemmae and stratified *A. thaliana* seeds were grown for 21 days on solid ½ Gamborg media (Section 3.3.14.) supplemented with varying glyphosate doses. Plants were then imaged using a Berthold Nightowl II LB 983 *In Vivo* Imaging System (Berthold, Bad Wildbad, Germany) after 120 s of white light exposure. The system detects chlorophyll autofluorescence at 560 nm, and the indiGO™ software package (Berthold, Bad Wildbad,

Germany) was used to determine the area of living auto-fluorescing tissue. Dose-response curves were generated by fitting a four-parameter log-logistic equation to the data using GraphPad Prism version 9 for Mac OS X. IC50 values were interpolated from the curves.

3.3.9. Site-directed mutagenesis

To introduce mutations into the overexpression vectors, I used the Agilent QuikChange II site-directed mutagenesis kit. This kit employs a two-step PCR-based method with template plasmids containing the wild-type *A. thaliana* and *M. polymorpha* EPSPS genes. Primers with desired mutations were designed using Agilent's primer design software, and PCR was performed according to the manufacturer's instructions. Following PCR, DpnI endonuclease digested the parental DNA template, leaving only the mutated fragments. Mutated plasmids were transformed into *E. coli* cells, selected on spectromycin-containing LB plates, and colonies with mutated plasmids were chosen, cultured, and their plasmid DNA was extracted using a purification kit. Sanger sequencing confirmed the desired mutations. These plasmids with mutations were then transformed into *Agrobacterium* for *A. thaliana* transformation (as described in Section 3.3.6).

3.3.10. Crude-EPSPS isolation

Protein extracts containing EPSPS were obtained from *A. thaliana*, *M. polymorpha*, *L. cruciata*, and *C. conicum*. Frozen leaf or thallus tissue (5 g) was ground into a fine powder, mixed with ice-cold extraction buffer (1g:5ml) containing 1% polyvinylpyrrolidone (PVP) and 5 mM fresh β-mercaptoethanol, and homogenised for 5 minutes at 4 ℃. After centrifugation, the supernatant was strained through a 50 µM cell strainer. Ammonium sulfate was added to reach 45%, followed by another centrifugation and pellet collection. Subsequently, ammonium sulfate was added to the supernatant until 70%, followed by centrifugation and pellet collection. This pellet was dissolved in 5-10 ml of extraction buffer with fresh 5 mM β-mercaptoethanol. Dialysis was carried out overnight using a Thermo-Scientific Slide-A-Lyzer cassette-based dialysis kit (10,000 MWCO) in 4 L dialysis buffer. The enzyme solution was concentrated using a Thermo-Scientific Pierce protein concentrator (10,000 MWCO). The resulting protein was used for enzyme assays to create a glyphosate dose-response curve, as described in Section 3.3.12.

3.3.11. EPSPS expression and purification

Six EPSPS proteins from *E. coli*, *A. tumefaciens* CP4 strain, *Z. mays* WT, *Z. mays* with TIPS mutation, *M. polymorpha*, and *A. thaliana* were expressed and purified. Expression constructs were transformed into *E. coli* strain BI21(DE3), cultured in liquid LB media with kanamycin at 37 ℃ overnight. Cultures were then diluted 1:250 in 2 L TB media with kanamycin and 1.5% w/v lactose and incubated with shaking (220 rpm) at 25 ℃ for 24 h. After centrifugation at 21130 rcf at 4 ℃ for 15 mins, pellets were resuspended in 150 ml of Ni A buffer with Benzonase and protease inhibitors, followed by sonication (19 mm probe, 5 min: Pulse on 1s, pulse off 2s, 60% amplitude). The intracellular soluble fraction was isolated by centrifugation at 20000 g for 30 mins at 4 ℃. His-tagged proteins were purified using immobilized metal affinity chromatography (IMAC) with 5 ml HisTrap FF columns, followed by 3C protease treatment and dialysis. Reverse IMAC was performed to remove the His-tag, and the tag-free protein was obtained in the flowthrough. Samples were concentrated to 1.5 ml using vivaspin 20 (MWCO 10 kDa) and subjected to size exclusion chromatography on a HiLoad 16/60 Superdex200 pg column equilibrated with SEC buffer at a flow rate of 1 ml/min.

3.3.12. EPSPS enzyme assays

Purified EPSPS enzyme activity and kinetics, as well as those from crude extracts, were assessed following an adapted protocol based on Priestman et al. (2005) and Funke et al. (2006). Enzymatic activities were determined by measuring inorganic phosphate production in the forward reaction using S3P and PEP as substrates in 96-well plates with a BioTek Synergy H1 plate reader. The quantification relied on a malachite green phosphate assay kit, and absorbance changes at 650 nm were recorded. Enzyme activity was expressed as Units mg^{-1} (U mg^{-1}), representing micromoles of phosphate produced, per minute, per milligram of enzyme. Each 80 µl reaction mixture contained 50 mM HEPES (pH 7.5), 100 mM KCl, and 2 mM DTT. EPSPS kinetics concerning S3P and PEP were studied by holding one ligand at a constant saturation concentration (1 mM) while varying the other. Enzyme activity changes were measured and fitted to the Michaelis-Menten equation to calculate kinetic parameters. Glyphosate tolerance was assessed using S3P and PEP at saturation concentration alongside varying glyphosate doses. Enzymatic reactions were initiated with enzyme addition and

incubated for 30 minutes at 25 °C, followed by halting the reactions with malachite green, allowing for 30 minutes of colour development.

3.3.13. Primers

Primer	Sequence	Use
Marchantia_EPSPS_fwd	AACAGGTCTCAAACAATGGAGTCTCTGAGG	Amplification of MpEPSPS with BsaI overhangs from gDNA
Marchantia_EPSPS_rev	AACAGGTCTCAGCAGCTAGGCAGACGAAAATTG	Amplification of MpEPSPS with BsaI overhangs from gDNA
Arabidopsis_EPSPS_fwd	AACAGGTCTCAGCAGCTAGGCAGACGAAAATTG	Amplification of AtEPSPS with BsaI overhangs from gDNA
Arabidopsis_EPSPS_rev	AACAGGTCTCAGCAGTTAGTGCTTTGTGATTC	Amplification of AtEPSPS with BsaI overhangs from gDNA
MpEPSPS_S500V_1	GGATCTCTGATGATAATTTTCACGTCACCACAAGCAGCAAGAGA	Site-directed mutagenesis of MpEPSPS
MpEPSPS_S500V_2	TCTCTTGCTGCTTGTGGTGACGTGAAAATTATCATCAGAGATCC	Site-directed mutagenesis of MpEPSPS
AtEPSPS_V_493S_1	CGTTGATGGTGATTGGAGAATCAGCACAAGCTGCAAGAGAG	Site-directed mutagenesis of AtEPSPS
AtEPSPS_V_493S_2	CTCTCTTGCAGCTTGTGCTGATTCTCCAATCACCATCAACG	Site-directed mutagenesis of AtEPSPS
AtSDM_fwd	TACAGTGGAAGAAGGTTCAG	Genotyping AtEPSPS site-directed mutagenesis construct
AtSDM_rev	ATTGAACTTGACGAACGTTG	Genotyping AtEPSPS site-directed mutagenesis construct
MpSDM_fwd	ACGCAGCCATTGACACTTACG	Genotyping MpEPSPS site-directed mutagenesis construct
MpSDM_rev	TTGTCGAAACCGATGATACG	Genotyping MpEPSPS site-directed mutagenesis construct

3.3.14. Buffers and Media

½ Gamborg medium

Component	**Final Concentration**
Gamborg B5 medium basal salt mixture	1.5 g/l (w/v)
MES	0.5 g/l (w/v)
Sucrose	1 % (w/v)
Agar (Dechefa)	0.8 % (w/v)

pH adjusted to 5.5 using 1M KOH

MS medium

Component	**Final concentration**
MS salt (no vitamins)	2.2 g/l (w/v)
MES	0.5 g/l (w/v)
Agar (Duchefa)	0.8 % (w/v)

pH adjusted to 5.8 using 1M KOH

Extraction buffer

Component	**Final concentration**
MOPS	100 mM
EDTA	5 mM
Glycerin	10 % (w/v)
KCl	50 mM
Benzamidine	0.5 mM
Pepstatin	7.3 µM
Trypsin Inhibitor	2.5 mg/l (w/v)
Leupeptin	4.2 µM

pH adjusted to 7 using 1M KOH, stored at 4°C

Dialysis buffer

Component	**Final concentration**
MOPS	10 mM
Glycerin	5 % (w/v)
EDTA	0.5 mM
B-mercaptoethanol	2.5 mM

3.4. Results

3.4.1. *M. polymorpha* plants are more tolerant to application of glyphosate than *A. thaliana* plants

To determine the growth response of wild-type *A. thaliana* and *M. polymorpha* plants to glyphosate treatment, a whole-plant dose response experiment was conducted (Fig. 3.2). Tak-1 was used as the wild-type *M. polymorpha* line, and Col-0 as the wild-type *A. thaliana* line. *A. thaliana* and *M. polymorpha* plants were grown for 21 days on media supplemented with varying glyphosate concentrations before plants were imaged. Growth reductions were observed in both species compared to plants grown on un-supplemented media, however, *M. polymorpha* was many times more tolerant to glyphosate than *A. thaliana* (Fig. 3.2 A). Glyphosate begins to inhibit *A. thaliana* growth, measured by rosette surface area, at a dose between 0.5 and 1 µM of glyphosate, with 1 µM offering almost complete control of *A. thaliana*. Conversely, there was no inhibition of *M. polymorpha* growth, measured by thallus surface area, up to a dose of 5 µM of glyphosate. Between 5 µM and 10 µM of glyphosate, *M. polymorpha* growth was inhibited, but even at 10 µM of glyphosate, full control of *M. polymorpha* was not achieved.

A four-parameter log-logistic equation was fitted to the dose response data, and from this, the glyphosate dose at which growth was reduce by half (IC_{50}) was calculated. This offers a quantified value of the magnitude of tolerance of each species (Fig. 3.2 B). The IC_{50} of glyphosate for *M. polymorpha* is 7.3560 µM (SE = 2.268) whereas the IC50 of glyphosate for *A. thaliana* is more than 12 times lower at 0.5978 µM (SE = 0.02647). These IC_{50} values are significantly different (Welch's unpaired t-test, p = 0.0155) and indicate a high level of glyphosate tolerance in *M. polymorpha* compared to *A. thaliana*.

The lethal dose of glyphosate is also higher in *M. polymorpha* compared to *A. thaliana*. Plant area was measured from chlorophyll autofluorescence. Concentrations where autofluorescence was no longer detected were considered lethal. A dose of 5 µM of glyphosate killed all but two *A. thaliana* plants, and these were very small (mean area = 0.8 mm^2) (Fig. 3.2 C). At higher doses of glyphosate there was no living tissue measured, so the lethal dose of glyphosate to *A. thaliana* is conservatively estimated as 10 µM, although it is

likely lower. Glyphosate has no discernible impact on *M. polymorpha* chlorophyll autofluorescence at doses up to and including 5 µM. At higher doses, there is inhibition of *M. polymorpha* by glyphosate, shown by a reduction in the area of auto-fluorescing tissue in *M. polymorpha* compared to in control conditions. However, even at 100 µM of glyphosate, 2 plants were still alive, although small (mean area = 1.1 mm^2). Therefore, the lethal dose of glyphosate to *M. polymorpha* is greater than 100 µM. *M. polymorpha* is therefore many times more tolerant to glyphosate than *A. thaliana*. To determine if this is due to known resistance-causing amino acid or structural differences in *M. polymorpha* EPSPS, a phylogenetic and structural analysis of EPSPS proteins was carried out.

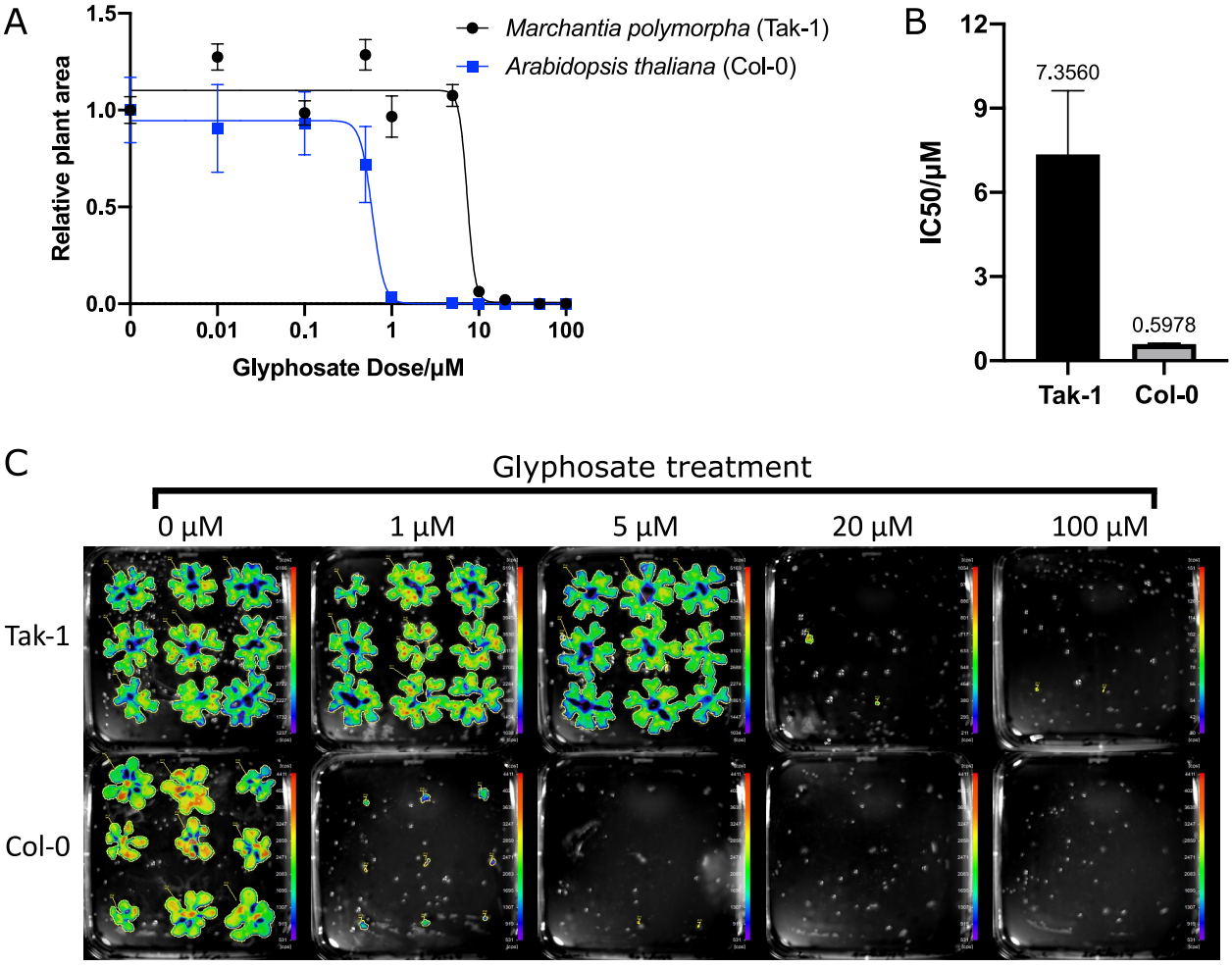


Fig. 3.2. A glyphosate dose response experiment reveals *M. polymorpha* to be more tolerant to glyphosate than *A. thaliana*

A. Dose response curve of the plant area of *A. thaliana* and *M. polymorpha* plants grown for 21 days on glyphosate-treated plates. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE.

B. The IC_{50} values of glyphosate for *A. thaliana* (Col-0) and *M. polymorpha* (Tak-1) calculated from the regression curve fitted to the dose response data. Error bars are ±SE. The reported IC_{50} values are significantly different (Welch's unpaired t-test, p = 0.0155).

C. Images of chlorophyll auto-fluorescing area of *A. thaliana* (Col-0) and *M. polymorpha* (Tak-1) plants on 5 selected doses of glyphosate.

3.4.2. The EPSPS protein sequence and structure is highly conserved among land plants, including *M. polymorpha*

Glyphosate resistance in many weeds is attributed to amino acid changes causing structural alterations in EPSPS proteins, known as target-site resistance (TSR) (Heap, 2023). To assess whether glyphosate tolerance in *M. polymorpha* is due to TSR, I aligned EPSPS enzyme amino acid sequences from 27 land plants. This alignment served as the basis for constructing a phylogenetic tree using maximum likelihood statistics (Fig. 3.3 A). Alignments were manually trimmed to remove extensive gaps, preserving mostly conserved sequences. *K. nitens*, a streptophyte alga, was chosen as the outgroup, and the phylogenetic tree was rooted at this node. The tree topology reveals a monophyletic clade of land plant EPSPSs. Within this clade, the three eudicot EPSPS sequences analysed form a monophyletic clade, sister to the monophyletic clade formed by two EPSPS sequences from monocots. Sister to the eudicot and monocot EPSPS clade is the sequence from *Amborella trichopoda*, forming a well-supported monophyletic angiosperm EPSPS clade. The single gymnosperm EPSPS sequence in this analysis, from *Gingko biloba*, is sister to the angiosperm EPSPS sequences, forming a monophyletic clade of seed plant EPSPS sequences. Two fern EPSPS sequences from *Ceratopteris richardii* and *Salvinia cucullata* create a monophyletic clade, sister to the seed plant EPSPS sequences, resulting in a large monophyletic clade of vascular plant EPSPS sequences. However, the EPSPS sequence from the lycophyte *Selaginella moellendorffii* is grouped with EPSPS sequences from hornworts, quillworts, and clubmosses, outside the vascular plant clade. Additionally, EPSPS sequences from five liverwort species, including *M. polymorpha*, form a monophyletic clade sister to the large clade composed of vascular plants, hornworts, quillworts, and clubmosses. Lastly, EPSPS sequences from five moss species analysed create a monophyletic group, positioned as the sister clade to EPSPS sequences from all other land plants. The distinct topology of the EPSPS phylogenetic tree, compared to widely accepted land plant phylogenies, is likely a consequence of the high sequence similarity among land plant EPSPSs. A phylogeny of EPSPS sequences encompassing the Archaeplastida is provided in Supplementary Fig. S3.1.

The EPSPS sequences analysed were highly conserved, with high percentage identities between the sequences from land plant species. This has likely made the topology of the

tree difficult to resolve. This high level of sequence similarity is especially so around the active site of the enzyme and the GTAMRP domain (Fig. 3.3 A). There are no amino acid sequence differences in this domain across all the land plants. In the field, most TSR cases are caused by at least one amino acid change in this domain. *M. polymorpha* also has the GTAMRP sequence at this position, and so a change in amino acid sequence in this domain is not the cause of glyphosate tolerance in *M. polymorpha*.

However, amino acid differences in alternative parts of the EPSPS sequence of *M. polymorpha* may cause structural changes sufficient to confer glyphosate tolerance. In order to search for structural differences in the EPSPS enzyme of *M. polymorpha* compared to *A. thaliana*, homology models of the EPSPS enzymes from both species were generated and compared using ChimeraX (Fig. 3.3 B). The 3D structure of homology models of the two proteins were almost exactly identical. The sequence alignment score for AtEPSPS and MpEPSPS was 1793.3, with an RMSD of 0.223 Å for 421 pruned atom pairs (0.60 Å for all 439 pairs). Therefore, large structural differences in the EPSPS protein are unlikely to be the cause of glyphosate tolerance in *M. polymorpha*. However, it is formally possible that subtle differences in the properties of the enzymes, due to amino acid differences, may hinder glyphosate binding to *M. polymorpha* EPSPS and are not captured in this analysis. In order to test this, an *in silico* docking analysis was performed to test if glyphosate is modelled to bind to the homology models of *M. polymorpha* EPSPS.

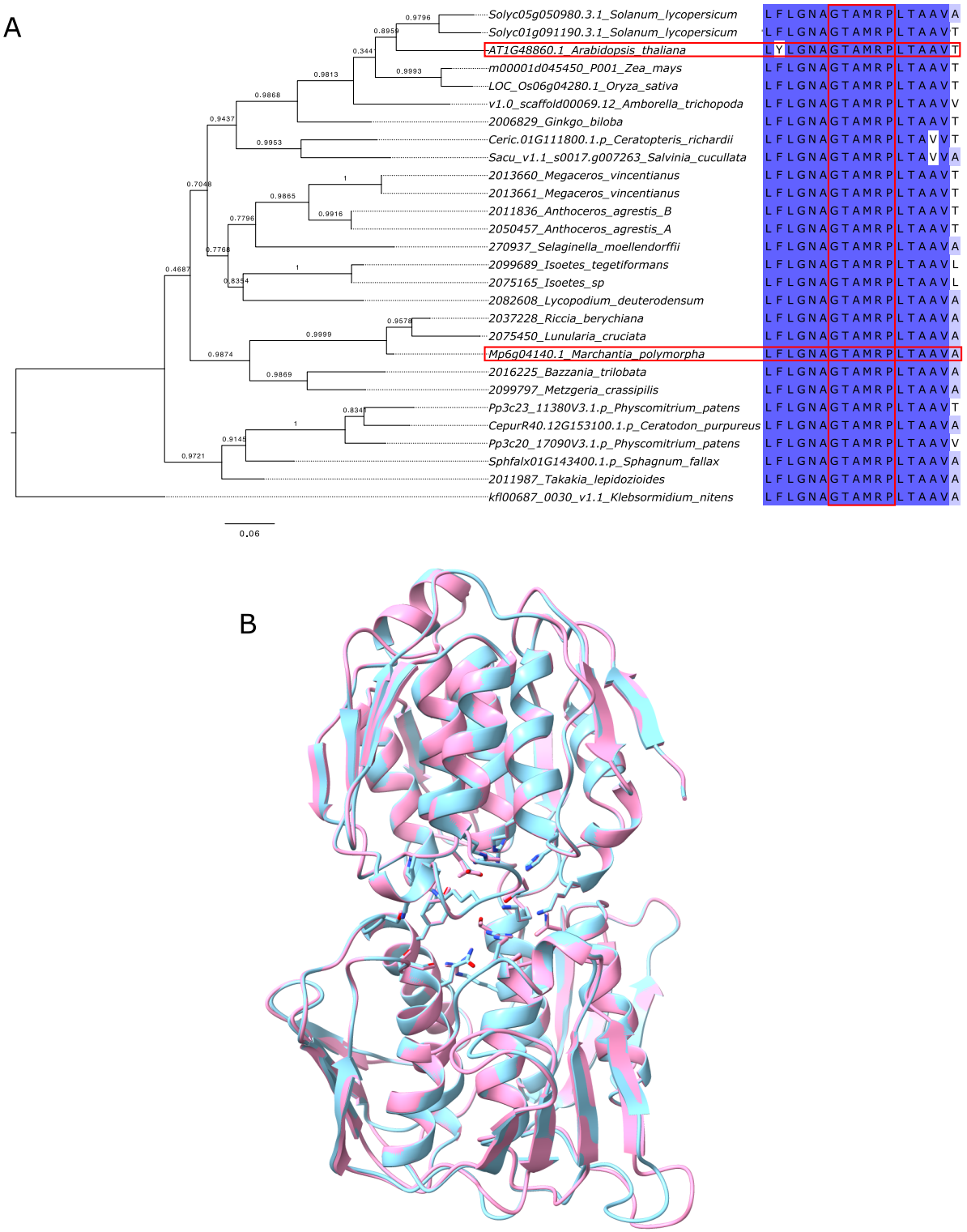


Fig. 3.3. Comparative analysis of *M. polymorpha* EPSPS sequence with other land plant EPSPSs sequences and *M. polymorpha* EPSPS structure with *A. thaliana* EPSPS structure

A. Maximum likelihood tree of land plant EPSPS amino acid sequences rooted to the EPSPS sequence from *K. nitens*. Phylogenetic tree was constructed using PHyML 3.0 (Guindon et al., 2010) using an estimated gamma distribution model and the WAG+G+I model of amino acid substitution. Support values are given from a Chi^2-based aLRT test. Protein sequences were aligned using MAFFT with the L-INS-I algorithm. The alignment of the GTAMRP domain is given.

B. Overlayed homology models of *A. thaliana* EPSPS (pink) and *M. polymorpha* EPSPS (blue) made in SwissModel based on the crystal structure of the EPSPS from *Vibrio cholerae* (PDB D: 3nvs, unpublished).

3.4.3. Glyphosate is modelled to bind to the EPSPS of *M. polymorpha* in the same confirmation as to a glyphosate-resistant EPSPS enzyme

To model how and if glyphosate binds to *M. polymorpha* EPSPS, an *in silico* molecular docking analysis was conducted. Additionally, I performed the same analysis on homology models of five other EPSPS enzymes for comparison. These were from two glyphosate sensitive plants, *A. thaliana* and *Z. mays*; a glyphosate sensitive bacterium, *E. coli*; and two resistant EPSPS enzymes used in glyphosate tolerant crops, *Z. mays* EPSPS with the glyphosate-resistance-conferring TIPS mutation, and the EPSPS from the glyphosate-resistant CP4 strain of *A. tumefaciens*. The amino acid sequence of each EPSPS was collected from Uniprot and the homology model constructed in SwissModel using the crystal structure of the EPSPS from *Vibrio cholerae* as a template (PDB D: 3nvs, unpublished). Molecular docking of glyphosate along with shikimate-3-phosphate (S3P) was then carried out using SeeSar. The 2D interaction maps of the glyphosate molecules bound to each EPSPS model are presented in Fig. 3.4.

Glyphosate was predicted to bind to the binding site of all 6 of the tested EPSPS enzyme models. The EPSPS enzymes from *A. thaliana*, *E. coli* and *Z. mays* are known to be sensitive to glyphosate, and glyphosate is modelled to bind to these enzymes in the same conformation (Fig. 3.4 A-C). Glyphosate is predicted to interact with different amino acids and form different types of bonds with these three EPSPS enzymes, however the glyphosate molecule itself does not change shape and forms numerous interactions, including at least one covalent bond. Conversely, no covalent bonds are predicted to form between glyphosate and the two confirmed resistant EPSPS enzymes (from *A. tumefaciens* CP4 and *Z. mays* TIPS) (Fig. 3.4 D and E). Furthermore, the structure of glyphosate changes upon binding with these resistant enzymes, in both cases losing the more linear conformation of its natural state and acquiring a compact conformation. In the glyphosate molecule, this change is due to a rotation of the bond between the C2 and C3 carbon atoms in the case of the *A. tumefaciens* CP4 enzyme and a rotation of the bonds between the C1 and P1 atoms in the case of the *Z. mays* TIPS enzyme. No covalent bonds are predicted to form between the EPSPS of *M. polymorpha* and glyphosate, as with the resistant EPSPS enzyme models. Furthermore, glyphosate is predicted to bind to *M. polymorpha* EPSPS in the same compact conformation as in the model of glyphosate bound to the *A. tumefaciens* CP4 enzyme, also

with a rotation of the bond between the C2 and C3 carbons. This similarity in binding prediction is consistent with glyphosate not binding efficiently to *M. polymorpha* EPSPS causing it to be glyphosate-tolerant. In order to test this hypothesis that *M. polymorpha* has a glyphosate-tolerant EPSPS, I carried out an *in vitro* dose response assay on EPSPS enzymes from various sources, including *M. polymorpha*.

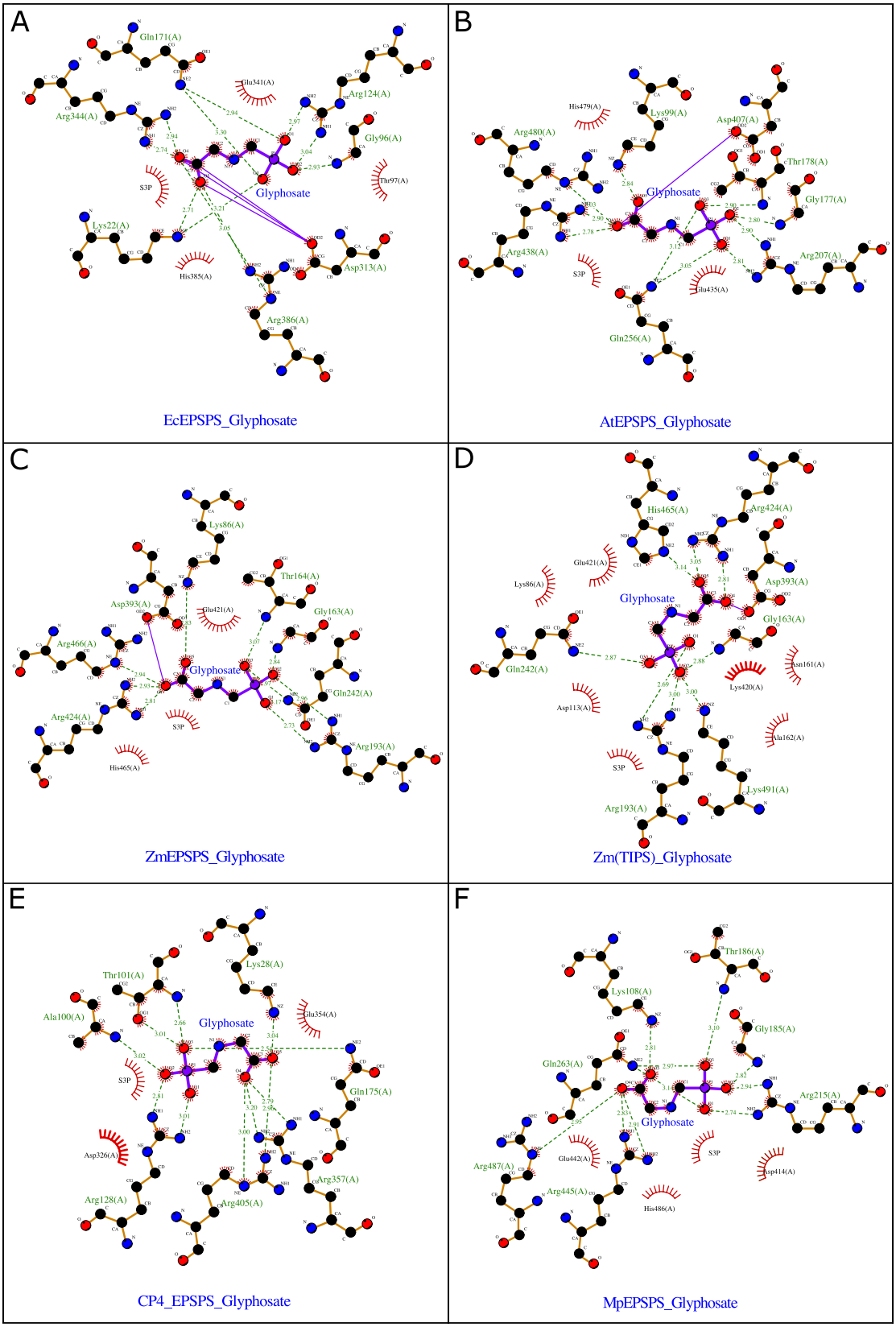


Fig. 3.4. 2D interaction maps of molecular docking results of EPSPS enzymes bound to glyphosate

2D interaction maps produced in LigPlot+ v.2.2 of homology models of EPSPS proteins from *E. coli* (A), *A. thaliana* (B), *Z. mays* (C), *Z. mays* with TIPS mutation (D), *A. tumefaciens* CP4 strain (E) and *M. polymorpha* (F), bound with glyphosate, after performing a molecular docking analysis in SeeSar. Purple bond lines represent covalent bonds between the protein and ligand, brown lines represent non-ligand bonds, green dashed lines represent hydrogen bonds, red crescents represent non-ligand residues involved in hydrophobic contacts and red dashes around atoms represent corresponding atoms involved in hydrophobic contacts.

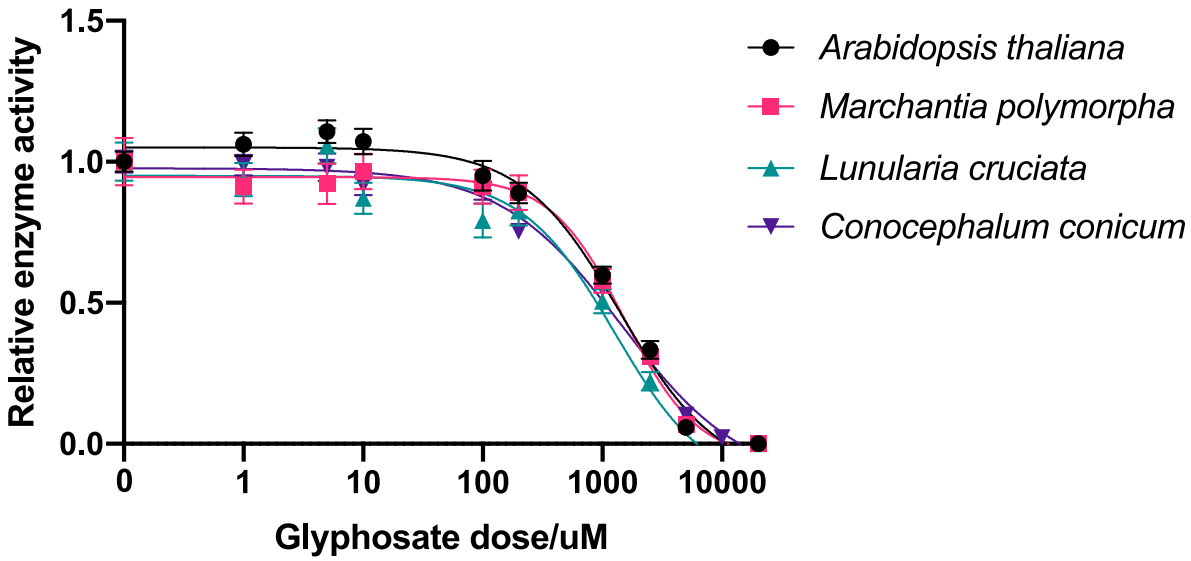
3.4.4. *M. polymorpha* crude protein extract containing EPSPS is not more tolerant to glyphosate than crude protein extracts containing EPSPS from other land plants

My molecular docking analysis predicts glyphosate to bind to *M. polymorpha* EPSPS in a similar way as to the EPSPS from *A. tumefaciens* CP4, which is resistant to glyphosate. In order to test if *M. polymorpha* EPSPS is tolerant to glyphosate, protein extract containing EPSPS was taken from *M. polymorpha,* and the enzymatic activity of this extract measured at varying doses of glyphosate (Fig. 3.5). Activity was measured by measuring the accumulation of inorganic phosphate, one of the products of the reaction which EPSPS catalyses. Protein extracts containing EPSPS were taken from *A. thaliana*, *M. polymorpha*, *Lunularia cruciata* and *Conocephalum conicum*. Images of the liverwort species can be found in Supplementary Fig. S3.5. *A. thaliana* is sensitive to glyphosate and so its EPSPS is sensitive to glyphosate. *M. polymorpha* is tolerant to glyphosate, and my molecular docking analysis predicts *M. polymorpha* EPSPS to be tolerant to glyphosate. I tested the activity of *A. thaliana* and *M. polymorpha* EPSPS enzymes extracted from plant tissue at different concentrations of glyphosate to determine if there was a difference in tolerance to glyphosate of these two proteins. *L. cruciata* and *C. conicum* are two other thalloid liverwort species and so the EPSPSs from these species were included to test if a glyphosate-tolerant EPSPS was specific to *M. polymorpha* or to liverworts generally.

Activity of the crude extracts at constant concentrations of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) (1 mM of each) were measured at 11 glyphosate concentrations between 0 and 20 mM. The relative enzyme activity compared to control was plotted and a four-parameter log-logistic regression fitted to the data (Fig. 3.5 A). The relative activity of protein extracts from the 4 species all showed a similar dose response relationship to glyphosate treatment. The activity of the protein extracts tested begin to be reduced at a dose of approximately 0.1-1 mM. Activity of all the extracts is completely arrested at a dose of approximately 5 mM. This is consistent with the EPSPS enzymes from all these species being sensitive to glyphosate. IC_{50} values for glyphosate of each of the enzyme extracts was interpolated from the dose-response curves and are presented in Fig. 3.5 B. The protein extract from *A. thaliana*, which is known to have a glyphosate sensitive EPSPS enzyme, has the lowest IC_{50} of 1377 µM glyphosate (SE = 169.9). The IC_{50} values obtained from protein

extracts from the liverwort species are all higher IC_{50} than that of *A. thaliana*, with *L. cruciata* having the highest at 1583 µM of glyphosate (SE = 256.0). The glyphosate IC_{50} calculated for *M. polymorpha* crude protein extract is 1553 µM (SE = 121.9) and that from *C. concicum* is 1460 µM (SE = 185.3). However, there is no significant differences between any of these IC_{50} values (one-way ANOVA, p = 0.9511). These data demonstrate that the protein extracts with EPSPS enzymes from these four species are all equally sensitive to glyphosate. This is consistent with *M. polymorpha* EPSPS not being glyphosate-tolerant. However, these are crude protein extracts and there is also no extract containing a known resistant EPSPS for comparison. In order to test more accurately the sensitivity of the *M. polymorpha* EPSPS to glyphosate, the dose-response relationship to glyphosate of purified *M. polymorpha* EPSPS and 5 other EPSPSs was measured.

A



B

Crude EPSPS source	*Arabidopsis thaliana*	*Marchantia polymorpha*	*Lunularia cruciata*	*Conocephalum conicum*
IC_{50} (µM glyphosate)	1377 (169.9)	1553 (121.9)	1583 (256.0)	1460 (185.3)

Fig. 3.5. Glyphosate dose-response experiment on the activities of protein extracts containing EPSPS from four land plant species

A. Dose-response curves of relative activities of crude protein extracts containing EPSPS enzymes from *A. thaliana, M. polymorpha, L. cruciata* and *C. conicum*. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE.

B. IC_{50} values interpolated from the dose-response curve. Standard errors are in parentheses.

3.4.5. Purified *M. polymorpha* EPSPS is sensitive to glyphosate

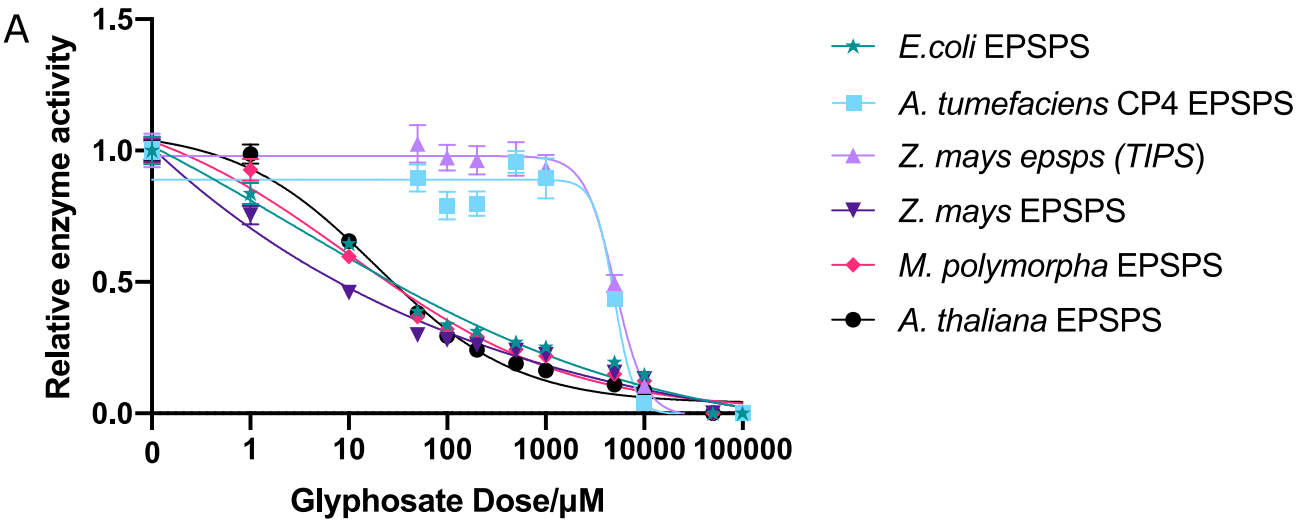
Dose response data from crude protein extracts suggest that *M. polymorpha* EPSPS is as sensitive to glyphosate as the EPSPS of a glyphosate sensitive plant, *A. thaliana*. In order to confirm this result, a dose response experiment using *M. polymorpha* EPSPS protein expressed in *E. coli* and purified was conducted. Three EPSPSs which were glyphosate sensitive were also expressed and purified, from *E. coli*, *A. thaliana* and *Z. mays.* Two EPSPS proteins that are used in glyphosate resistant crops and are resistant to glyphosate were also expressed and purified (Supplementary Fig. S3.6). These were the *Z. mays* EPSPS with the TIPS mutation introduced, and the EPSPS from the *A. tumefaciens* CP4 strain. Similarly to in Section 3.4.4, the activity of these enzymes was measured at differing concentrations of glyphosate while the concentrations of the two substrates of the reaction EPSPS catalyses, S3P and PEP, were kept constant at a saturating concentration (1 mM).

Activity was measured by the quantity of inorganic phosphate produced in the forward reaction. The relative activity of the enzyme compared to that at 0 µM glyphosate was plotted for each dose up to 100 mM. A four-parameter log-logistic regression was fitted to the data (Fig. 3.6 A). The dose-response relationship for the EPSPS enzymes from *E.coli*, *A. thaliana, Z. mays* (wild-type) and *M. polymorpha* to glyphosate was similar. Activity begins to decrease in these enzymes at a concentration of glyphosate as low as 1 µM. Based on the curves produced, the EPSPS of wild-type *Z. mays* appears to be the most sensitive to glyphosate. The curves of *A. thaliana*, *E. coli* and *M. polymorpha* EPSPSs are very similar, suggesting that *M. polymorpha* EPSPS is indeed as sensitive to glyphosate as the *A. thaliana* and *E. coli* EPSPSs. The activity of the two resistant EPSPS enzymes has an entirely different glyphosate dose-response relationship to the other EPSPS enzymes in this analysis. The activity of both the *Z. mays* EPSPS with the TIPS mutation and the EPSPS from *A. tumefaciens* CP4 is unchanged up to a dose of 1 mM. A dose of 10 mM is required to completely arrest the activity of these two enzymes. The dose-response curves of these glyphosate-resistant enzymes suggest the magnitude of their resistance is very similar.

The IC_{50} values of the enzymes were interpolated from the curves (Fig. 3.6 B). The values for the glyphosate-resistant EPSPSs are similar at 5055 and 5017 µM glyphosate for the *Z. mays* TIPS and *A. tumefaciens* CP4 EPSPS respectively (SEs = 228.0 and 235.6 respectively). The

difference in IC_{50} values of the two glyphosate-resistant EPSPSs is not statistically significant (one-way ANOVA, $p > 0.9999$). The IC_{50} values of each of these resistant EPSPS enzymes is significantly greater than each of the other four EPSPS enzymes (one-way ANOVA, $p < 0.0001$ for all comparisons). The IC_{50} values for glyphosate for the sensitive EPSPS enzymes were also not significantly different from each other (one-way ANOVA, $p > 0.95$). This is therefore strong evidence that the *M. polymorpha* EPSPS is as sensitive to glyphosate as the EPSPS enzymes from *A. thaliana, Z. mays* and *E. coli*. It can therefore be concluded that the *M. polymorpha* EPSPS is sensitive to glyphosate and therefore TSR is not the cause of glyphosate tolerance in *M. polymorpha*.

A



B

EPSPS source	*Escherichia coli*	*Agrobacterium tumefaciens CP4*	*Zea mays* (TIPS)	*Marchantia polymorpha*	*Zea mays*	*Arabidopsis thaliana*
IC_{50} (µM glyphosate)	22.19 (12.46)	5017 (235.6)	5055 (228.0)	9.566 (9.123)	2.021 (8.599)	14.49 (5.983)

Fig. 3.6. Glyphosate dose-response experiment on various purified EPSPS enzymes
A. Dose-response curves of relative activities of purified EPSPS enzymes from *A. thaliana, M. polymorpha, Z. mays* (wild-type), *Z. mays* (with TIPS mutation), *E. coli* and *A. tumefaciens* CP4 strain. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE.
B. IC_{50} values interpolated from the dose-response curve. Standard errors are in parentheses.

3.4.6. The kinetic properties of plant EPSPSs are similar to each other but different from bacterial EPSPS enzymes

The activities of the purified EPSPS enzymes used in the glyphosate dose-response experiment (Section 3.4.5) differed greatly. I hypothesised that EPSPS activity may influence glyphosate tolerance at the whole-organism level. To precisely assess the kinetic properties of these purified EPSPS proteins for their respective substrates, I measured their activity across varying concentrations of either S3P or PEP, while keeping the other substrate constant at a saturating concentration. Enzyme activity was determined by quantifying phosphate production during the forward EPSPS reaction, expressed as µmol phosphate produced/min of reaction time/mg of EPSP synthase (U mg^{-1}). Subsequently, I fitted the Michaelis-Menten equation to the data, and the resulting curves can be found in Supplementary Fig. S3.7 and S3.8. These curves were then used to calculate various kinetic parameters for each EPSPS enzyme, as presented in Table 3.1. Statistical differences between the kinetic parameters were tested using Welch's ANOVA and Dunnett's T3 multiple comparison test.

All tested EPSPS enzymes displayed expected saturation behaviour when specific activities were plotted against PEP concentration (Supplementary Fig. S3.7). The affinity of EPSPS for PEP, indicated by K_m, did not significantly differ among plant or bacterial EPSPS proteins. However, the Welch's ANOVA test indicated statistically significant differences in the calculated K_m values for EPSPS proteins ($p = 0.0003$). The bacterial enzymes, *E. coli* and *A. tumefaciens* CP4 EPSPS exhibited the highest PEP affinity, with low K_m values of 0.08338 and 0.08729 mM, respectively (Table 3.1). These values were not significantly different from each other, suggesting equal PEP affinity (adjusted p-value > 0.9999). In contrast, wild-type EPSPS enzymes from the plants *A. thaliana*, *M. polymorpha*, and *Z. mays* displayed higher K_m values, indicating lower PEP affinity. *A. thaliana* EPSPS had the lowest K_m at 0.1636 mM, followed by *M. polymorpha* at 0.1803 mM, and *Z. mays,* with the highest K_m and lowest PEP affinity at 0.2237 mM. However, these values were not significantly different, indicating that the plant EPSPS enzymes in this study shared the same PEP affinity (p-values > 0.94). The introduction of the TIPS mutation to *Z. mays* did not alter PEP affinity, as its K_m remained similar to that of wild-type plant EPSPS enzymes ($K_m = 0.1654$, adjusted p-values > 0.8). Only

Z. mays EPSPS (wild-type and TIPS) exhibited significantly different K_m values compared to other enzymes, particularly bacterial EPSPS enzymes (adjusted p-values < 0.05). In summary, differences in PEP affinity among EPSPS enzymes were not correlated with glyphosate tolerance.

Despite largely similar Michaelis-Menten constants calculated for each enzyme, the maximum velocities of the enzymes (V_{max}) for PEP varied significantly. V_{max} for each EPSPS enzyme differed significantly from every other enzyme tested (adjusted p-values < 0.05). *E. coli* had the highest V_{max} at 92.58 U mg^{-1}, followed by *A. tumefaciens* CP4 at 16.55 U mg^{-1}. Among plant EPSPSs, *A. thaliana* EPSPS exhibited the highest V_{max} at 26.81 U mg^{-1}. *M. polymorpha* and *Z. mays* EPSPS displayed more similar V_{max} values of 7.880 and 10.38 U mg^{-1}, respectively, though the difference remained significant (adjusted p-value = 0.0144). *M. polymorpha* EPSPS had the lowest V_{max} for PEP among wild-type plant EPSPS enzymes. The *Z. mays* EPSPS with the TIPS mutation showed a V_{max} approximately an order of magnitude lower than other plant EPSPS enzymes (0.6788 U mg^{-1}), indicating that the mutation significantly affected the enzyme's maximum velocity for this reaction. V_{max} values under varying S3P concentrations were similar and followed the same pattern as with varying PEP concentrations. In summary, differences in V_{max} for PEP and S3P were not correlated with glyphosate tolerance in the purified EPSPS enzymes.

The K_{cat} of the enzymes was also calculated from the Michaelis-Menten curves, and this value similarly varied among the proteins, mirroring the pattern observed for V_{max} values. K_{cat} represents the turnover rate of the enzyme. It describes the maximum number of substrate molecules that an enzyme can convert into product molecules per active site per unit of time, when the enzyme is fully saturated with the substrate. The calculated K_{cat} values for PEP were used to determine K_{cat}/K_m ratios, representing catalytic efficiency, with bacterial enzymes exhibiting higher values than plant enzymes. *E. coli* EPSPS had the highest catalytic efficiency, approximately an order of magnitude greater than any other enzyme (K_{cat}/K_m = 2.842x105 $M^{-1}s^{-1}$, adjusted p-values < 0.0006). The K_{cat}/K_m of the EPSPS enzymes from *A. thaliana* and *A. tumefaciens* CP4 had the next highest values and were not significantly different from each other (adjusted p-value > 0.9999), suggesting equal catalytic efficiency regarding PEP. *A. thaliana* EPSPS was the most efficient among plant enzymes,

although its K_{cat}/K_m was not significantly higher than that of *Z. mays* (wild-type) EPSPS (adjusted p-value = 0.0516). Among wild-type plant EPSPS enzymes, *M. polymorpha* EPSPS was the least efficient, with a K_{cat}/K_m of 1.399x103 $M^{-1}s^{-1}$, not significantly different from *Z. mays* (wild-type) EPSPS (adjusted p > 0.999) but significantly lower than that of *A. thaliana* EPSPS (adjusted p-value = 0.0490). The *Z. mays* EPSPS enzyme with the TIPS mutation exhibited the lowest catalytic efficiency for PEP, with a K_{cat}/K_m ratio of 8.210 $M^{-1}s^{-1}$, significantly lower than all other EPSPS enzymes tested (adjusted p-values < 0.05). In summary, differences in catalytic efficiency among the purified enzymes were not correlated with glyphosate tolerance.

The kinetic parameters for the EPSPS substrate S3P revealed similar patterns as with PEP. Bacterial EPSPS enzymes exhibited higher S3P affinity (lower K_m) and velocity compared to wild-type plant EPSPS enzymes. *A. tumefaciens* CP4 had the highest S3P affinity at 0.02887 mM, which was not significantly different from *A. thaliana* and wild-type *Z. mays* EPSPS (adjusted p-value > 0.05). However, this lack of statistical significance likely results from high variance in plant EPSPS K_m values. Despite its lower S3P affinity, *A. thaliana* EPSPS exhibited a V_{max} of 30.94 U mg^{-1}, about one-third of *E. coli* EPSPS's V_{max} at 95.91 U mg^{-1} (adjusted p-value < 0.0001). *A. tumefaciens* CP4 EPSPS had a lower V_{max} at 17.10 U mg^{-1}, significantly less than *A. thaliana* EPSPS (adjusted p-value = 0.0036). *M. polymorpha* displayed the lowest V_{max} for S3P among plant EPSPS enzymes at 7.248 U mg^{-1}, significantly lower than wild-type *Z. mays* EPSPS (V_{max} = 10.81 U mg^{-1})(adjusted p-value = 0.0028). The *Z. mays* EPSPS with the double TIPS mutation had a much lower V_{max} for S3P (V_{max} = 0.7332 U mg^{-1}), indicating the mutation greatly affects S3P utilization. In summary, V_{max} differences for S3P were not correlated with glyphosate tolerance.

K_{cat} values for S3P followed the V_{max} pattern, with *E. coli* EPSPS being the most efficient. *A. tumefaciens* CP4 EPSPS was the second most efficient but had lower efficiency than *E. coli* (adjusted p-value = 0.0015). Among wild-type plant EPSPS enzymes, *A. thaliana* EPSPS was the most efficient, *M. polymorpha* the least, while *Z. mays* EPSPS had intermediate efficiency. However, wild-type plant EPSPS enzymes had similar catalytic efficiencies for S3P (adjusted p-values > 0.11). The *Z. mays* EPSPS with the TIPS mutation had the lowest catalytic efficiency for S3P, significantly lower than all other EPSPS enzymes tested (adjusted

p-values < 0.05). In summary, catalytic efficiency differences were not correlated with glyphosate tolerance.

In conclusion, these experiments revealed significant differences in some kinetic properties of the tested EPSPS enzymes, including wild-type enzymes from *A. thaliana*, *M. polymorpha*, and *Z. mays*. To assess their potential impact on glyphosate tolerance at the whole-plant level, I expressed *A. thaliana* and *M. polymorpha* EPSPS genes in *A. thaliana*.

	K_m (PEP)	V_{max} (PEP)	K_{cat} (PEP)	K_{cat}/K_m (PEP)	K_m (S3P)	V_{max} (S3P)	K_{cat} (S3P)	K_{cat}/K_m (S3P)	IC_{50}
	mM	U mg^{-1}	s^{-1}	$M^{-1}s^{-1}$	mM	U mg^{-1}	min^{-1}	s^{-1}	mM
E. coli	0.08338 (0.008858)	92.58 (2.367)	23.70 (0.6059)	$2.842x10^5$ ($3.1059x10^4$)	0.08651 (0.009585)	95.91 (2.582)	24.55 (0.6611)	$2.838x10^5$ ($3.236x10^4$)	22.19 (12.46)
A. tumefaciens (CP4)	0.08729 (0.01212)	16.55 (0.5598)	1.059 (0.03583)	$1.213x10^4$ ($1.7338x10^3$)	0.02887 (0.002354)	17.10 (0.2367)	1.094 (0.01515)	$3.789x10^4$ ($3.134x10^3$)	5017 (235.6)
Z. mays (TIPS)	0.1654 (0.01571)	0.6788 (0.01845)	0.001358 ($3.691x10^{-5}$)	8.210 (0.811)	0.09827 (0.008731)	0.7332 (0.01636)	0.001466 ($3.273x10^{-5}$)	14.918 (1.367)	5055 (228.0)
Z. mays	0.2237 (0.03393)	10.38 (0.4870)	0.3322 (0.01558)	$1.485x10^3$ ($2.358x10^2$)	0.1271 (0.02476)	10.81 (0.5646)	0.3460 (0.01807)	$2.722x10^3$ ($5.4904x10^2$)	2.021 (8.599)
M. polymorpha	0.1803 (0.02558)	7.880 (0.3271)	0.2522 (0.01047)	$1.399x10^3$ ($2.0677x10^2$)	0.1197 (0.01282)	7.248 (0.2050)	0.2320 (0.006561)	$1.938x10^3$ ($2.146x10^2$)	9.566 (9.123)
A. thaliana	0.1636 (0.03375)	26.81 (1.579)	1.716 (0.1011)	$1.0489x10^4$ ($2.225x10^3$)	0.1967 (0.04463)	30.94 (2.100)	1.980 (0.1344)	$1.0066x10^4$ ($2.384x10^3$)	14.49 (5.983)

Table 3.1. Summary table of the kinetic parameters of 6 EPSPS enzymes.
The IC_{50} values for glyphosate are also included. Standard error values are in parentheses.

3.4.7. Overexpression of the *A. thaliana EPSPS* gene leads to higher glyphosate resistance than overexpression of the *M. polymorpha EPSPS* gene

I have previously shown that *M. polymorpha* is more tolerant to glyphosate than *A. thaliana*. Although the amino acid sequences and structure of the glyphosate molecular target, EPSPS, are similar between the two species, the kinetic properties of the two enzymes differ significantly. To determine whether differences in the kinetic properties of *A. thaliana* and *M. polymorpha* EPSPS enzymes, outlined in Section 3.4.6, influence glyphosate resistance *in planta*, the genes encoding these two enzymes were expressed under the cauliflower mosaic virus (CaMV) 35S promoter in *A. thaliana.* The YFP protein was also expressed in *A. thaliana* under the same promoter as a control. Plants were then grown on varying doses of

glyphosate for 21 days before imaging and the resulting dose response curves are presented in Fig. 3.7 A. The growth responses to glyphosate were measured for 3 independent *A. thaliana* lines for each gene expressed and values pooled.

Growth reductions were observed when glyphosate doses were increased in all lines, however, there were large differences between *A. thaliana* plants expressing the *A. thaliana* EPSPS protein compared to those expressing the *M. polymorpha* EPSPS protein. The area of 35S:At*EPSPS* plants did not change up to a dose of 20 µM of glyphosate and then began to be reduced at a dose of 50 µM and beyond. Conversely, the area of 35S:Mp*EPSPS* plants began to be reduced with a glyphosate dose of 1 µM and 10 µM of glyphosate offered almost complete control, although was not completely lethal (Fig. 3.7 B). The 35S:Mp*EPSPS* plants were only slightly more resistant than the control 35S:YFP plants. A dose of 20 µM glyphosate was lethal to the control plants but not to 35S:Mp*EPSPS* plants, confirming their slightly increased resistance (Fig. 3.7 B).

IC_{50} values of glyphosate for each of the lines were interpolated from the four-parameter log-logistic regressions fitted to the data. These were 1.181 µM for 35S:YFP plants (SE ±0.1114), 40.01 µM for 35S:At*EPSPS* (SE ±4.553) and 2.229 µM for 35S:*MpEPSPS* (SE ±0.2955). The IC_{50} values for each 35S:At*EPSPS* and 35S:Mp*EPSPS* plants were compared to those for 35S:YFP plants using a Welch's t-test. The IC_{50} values of *A. thaliana* lines overexpressing both the *A. thaliana* and *M. polymorpha EPSPS* gene were significantly higher than that of *A. thaliana* lines expressing YFP ($p < 0.0001$ and $p = 0.0010$ respectively). Overexpressing both the *A. thaliana* and *M. polymorpha EPSPS* genes in *A. thaliana* increases glyphosate resistance, however, this effect is much higher with overexpression of the endogenous *A. thaliana EPSPS* gene. The reason for this difference is unknown. To determine what factors could lead to differences in kinetic properties and glyphosate resistance when expressed in *A. thaliana*, I carried out further *in silico* analyses of the *A. thaliana* and *M. polymorpha* EPSPS enzymes.

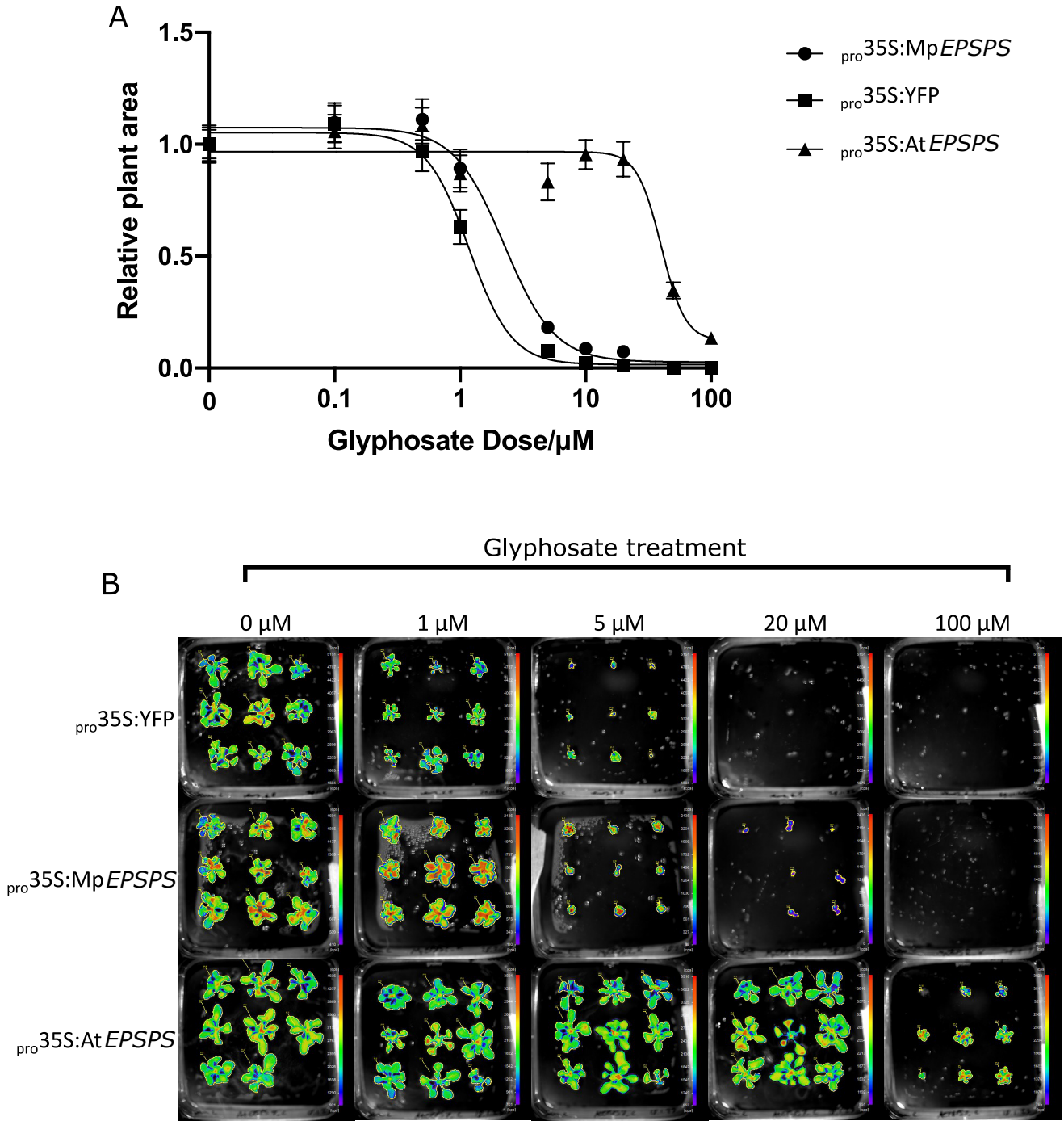


Fig. 3.7. A glyphosate dose response experiment reveals *A. thaliana* plants overexpressing the endogenous *EPSPS* are more resistant than those overexpressing the *M. polymorpha EPSPS*

A. Dose response curve of the plant area of *A. thaliana* expressing 35S:YFP, 35S:At*EPSPS* and 35S:Mp*EPSPS*. Plants were grown for 21 days on glyphosate-treated plates. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE. IC_{50} values for 35S:At*EPSPS* and 35S:Mp*EPSPS* plants are both greater than that for 35S:YFP plants ($p < 0.0001$ and $p = 0.0010$ respectively).

B. Images of chlorophyll auto-fluorescing area of *A. thaliana* plants on 5 selected doses of glyphosate.

3.4.8. *In silico* analysis of *A. thaliana* and *M. polymorpha* reveals a residue in *M. polymorpha* that may lead to differences in kinetic parameters

The EPSPS enzymes from *A. thaliana* and *M. polymorpha* differ in both kinetic properties and the resistance they confer when overexpressed in *A. thaliana*. To identify the causes of these distinctions, an *in silico* analysis was conducted to examine sequence, structure, and glyphosate binding differences. Amino acid sequence alignment revealed a distinct region around the 500th amino acid of *M. polymorpha* (Fig. 3.8 A). Further examination of homology models of both enzymes, presented in Section 3.4.2, indicated that the residue at position 500 in *M. polymorpha* (serine) connects two alpha helices, each with at least one residue in the active site (Fig. 3.8 B). These active site residues include Lys 512, His 486, and Arg 487. No other region in the alignment with amino acid differences between *A. thaliana* and *M. polymorpha* EPSPSs has such a direct relation to the enzyme's active site.

To assess the influence of the residue at position 500 in *M. polymorpha* on the enzyme's binding site, an *in silico* molecular docking experiment was conducted on EPSPS homology models. These models included versions with either valine or threonine at this position (from *A. thaliana*, *Z. mays*, and *E. coli*), as well as the *M. polymorpha* EPSPS with a serine. Glyphosate was subjected to docking, and the resulting 2D interaction map was analysed using LigPlot+ v.2.2 (Fig. 3.8 C). The docking simulations revealed that glyphosate bound to all three EPSPS enzymes in a conformation consistent with the absence of serine at this position, representing the typical glyphosate conformation. In contrast, when serine was present at this position in *M. polymorpha*, glyphosate adopted a more compact conformation due to a rotation of the bond between the C2 and C3 carbon atoms.

To confirm that this effect was specifically due to the presence of serine and not another factor, homology models of *A. thaliana* and *M. polymorpha* EPSPS enzymes were created with residues exchanged at this position. Subsequent molecular docking analysis revealed that with valine replacing serine at position 500, glyphosate was predicted to bind to the *M. polymorpha* EPSPS in a longer, more linear conformation. Conversely, when valine at the equivalent position 493 in the *A. thaliana* EPSPS was exchanged for serine, glyphosate was predicted to adopt the more compact conformation, akin to its binding to wild-type *M. polymorpha* EPSPS. This difference indicated that the amino acid at this position indeed

influenced the modelled enzyme's active site, as glyphosate exhibited distinct binding conformations based on the presence or absence of serine.

To corroborate these findings *in planta*, corresponding mutations were introduced into the EPSPS genes of *A. thaliana* and *M. polymorpha*, resulting in the *A. thaliana* EPSPS having a serine at position 493 and the *M. polymorpha* EPSPS having a valine at that position. Subsequently, these genes were expressed in *A. thaliana*.

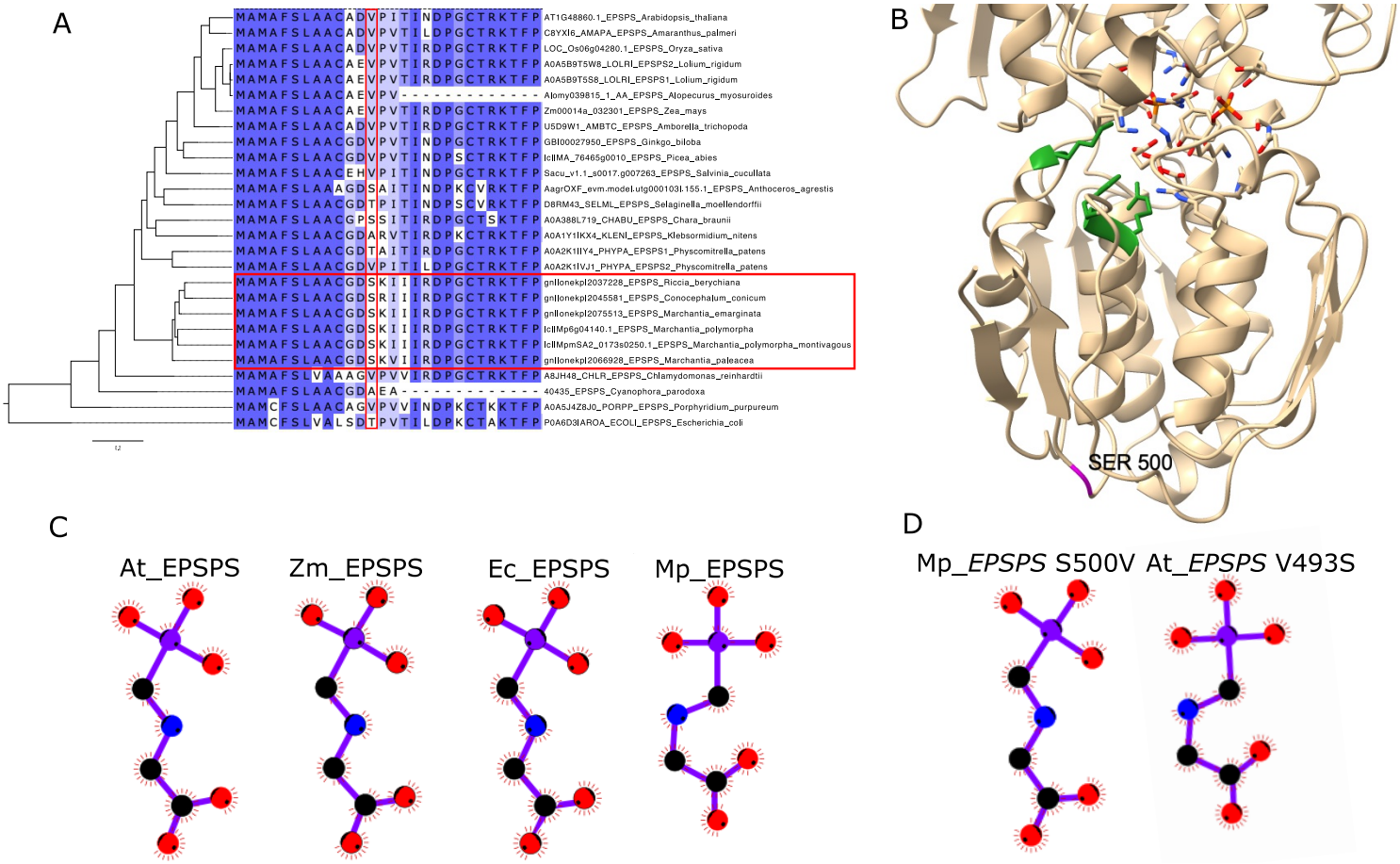


Fig. 3.8. *In silico* analysis of *A. thaliana* and *M. polymorpha* predicts a serine at position 500 in the *M. polymorpha* EPSPS to be important in the structure of the active site
A. Amino acid alignment of EPSPS sequences from various land plants rooted to the *E. coli* EPSPS sequence. Position 500 in *M. polymorpha* is highlighted, which has a serine in all liverwort species analysed. The maximum likelihood tree constructed from the whole alignment is given.
B. Homology model of the *M. polymorpha* EPSPS with glyphosate bound showing the position of the serine 500 (purple) and the residues of the active site connected by this residue (green).
C. Conformations of glyphosate molecules predicted to bind to wild-type EPSPS enzymes from *A. thaliana, Z. mays, E. coli* and *M. polymorpha*.
D. Conformations of glyphosate molecules predicted to bind to *M. polymorpha* EPSPS with a valine in place of serine at position 500 and *A. thaliana* EPSP with a serine in place of valine at position 493.

3.4.9. A serine to valine mutation at position 500 in the *M. polymorpha* EPSPS does not increase glyphosate sensitivity *in planta*

My *in silico* analysis in Section 3.4.8 predicted that a serine at position 500 in the *M. polymorpha* EPSPS influences the active site structure. To investigate its impact on glyphosate resistance *in planta*, I conducted site-directed mutagenesis of the *M. polymorpha* EPSPS gene, and overexpressed both the mutated and wild-type genes under the 35S promoter in *A. thaliana* and exposed the resulting plants to varying glyphosate doses. Three independent *A. thaliana* lines expressing each version of the *M. polymorpha* EPSPS were tested, and the data was pooled to generate dose-response curves (Fig. 3.9 A). IC_{50} values were interpolated from four-parameter log-logistic regressions, and the IC_{50} for plants expressing the wild-type EPSPS was slightly higher (2.229 µM of glyphosate) than those expressing the EPSPS with a valine substitution for the serine (1.558 µM of glyphosate). However, there was no difference in the IC_{50} values between the two protein versions (Welch's t-test, $p = 0.0697$). Therefore, substituting serine with valine at position 500 does not alter glyphosate sensitivity in plants expressing these proteins, as indicated by plant area images in Fig. 3.9 B.

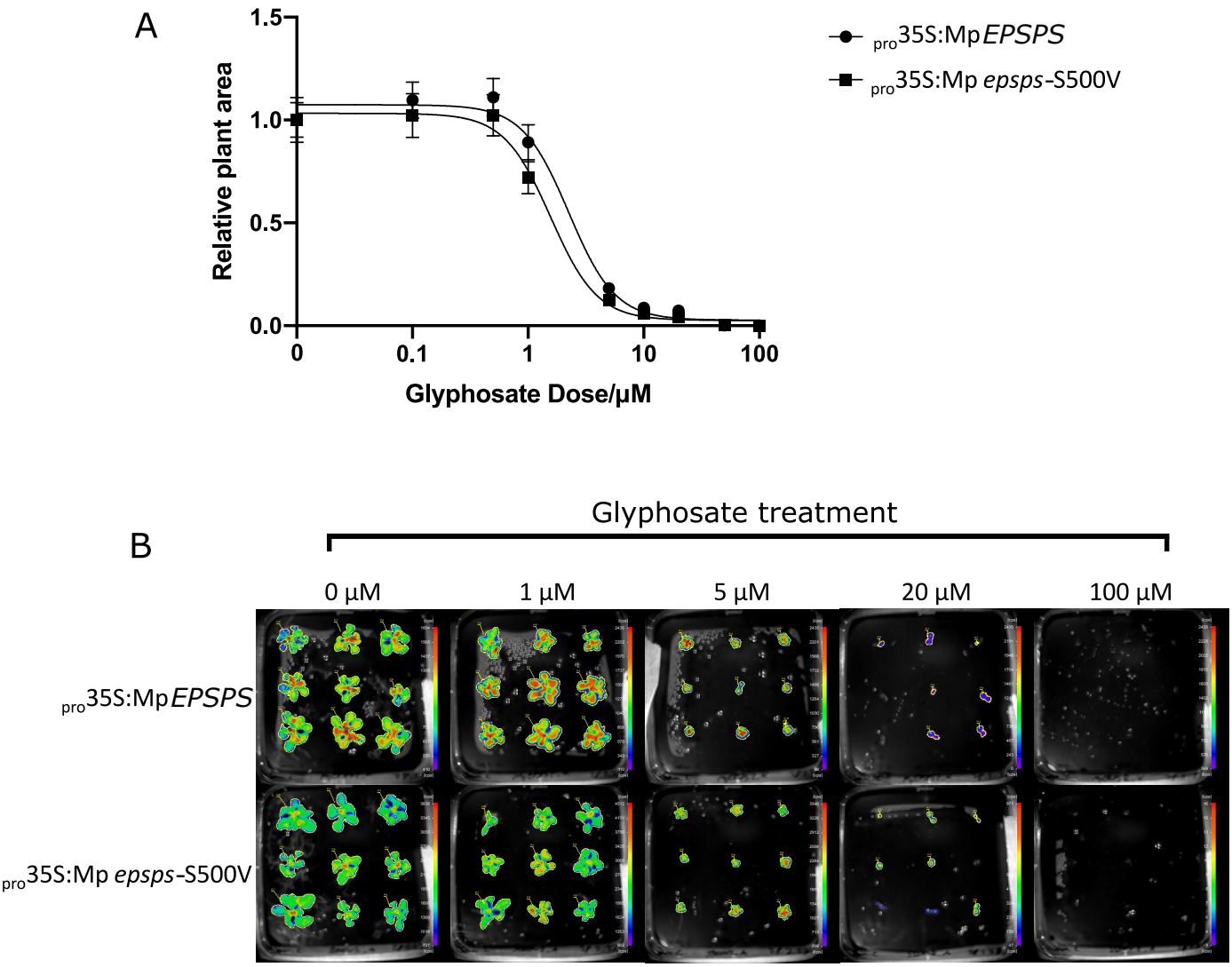


Fig. 3.9. A glyphosate dose response experiment reveals a valine for a serine at position 500 in *M. polymorpha* EPSPS does not increase glyphosate sensitivity *in planta*

A. A. Dose response curve of the plant area of *A. thaliana* expressing 35S:Mp epsps and 35S:Mp *epsps* S500V. Plants were grown for 21 days on glyphosate-treated plates. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE. IC_{50} values for 35S:Mp epsps and 35S:Mp *epsps* S500V are not significantly different (p = 0.0697).
B. Images of chlorophyll auto-fluorescing area of *A. thaliana* plants on 5 selected doses of glyphosate.

3.4.10. A valine to serine mutation at position 493 in the *A. thaliana* EPSPS does not increase glyphosate resistance *in planta*

A valine at position 493 in the *A. thaliana* EPSPS corresponds to a serine at position 500 in the *M. polymorpha* EPSPS. My *in silico* analysis (Section 3.4.8) suggests that valine may enhance glyphosate sensitivity, while serine provides tolerance. To test this, *A. thaliana* EPSPS was mutated, replacing valine with serine, and overexpressed in *A. thaliana*. Three independent lines for each version were subjected to glyphosate dose response experiments, with pooled data used to create dose-response curves (Fig. 3.10 A). IC_{50} values were interpolated from these curves. Plants overexpressing the wild-type version of the *A. thaliana* EPSPS were found to have an IC_{50} of 40.01 µM of glyphosate, while those plants expressing the EPSPS with a serine in place of the valine at position 493 had an IC_{50} of 23.17 µM. A Welch's t-test comparing these two IC_{50} values revealed a significant difference ($p = 0.0014$). However, this difference was not immediately evident in the measured plant areas (Fig. 3.10 B). These data are not consistent with the serine at this position increasing glyphosate resistance, instead showing that the serine substitution significantly decreases glyphosate resistance when comparing the IC_{50} values.

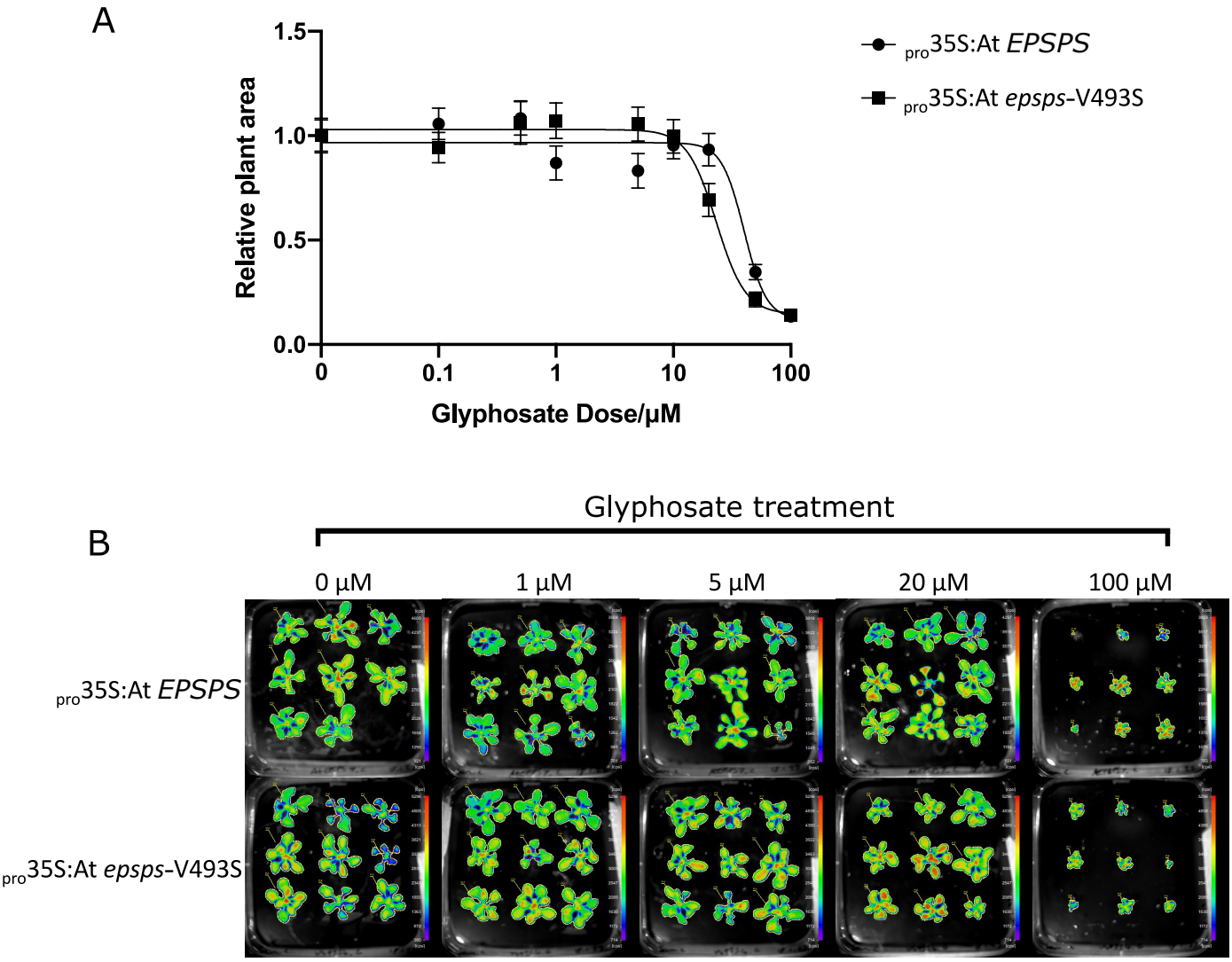


Fig. 3.10. A glyphosate dose response experiment reveals a serine for a valine at position 493 in *A. thaliana* EPSPS does not increase glyphosate resistance *in planta*

A. A. Dose response curve of the plant area of *A. thaliana* expressing 35S:At*EPSPS* and 35S:At*epsps* V493S. Plants were grown for 21 days on glyphosate-treated plates. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE. The IC_{50} value for plants expressing 35S:At*EPSPS* were significantly higher than those expressing 35S:At*epsps* V493S ($p = 0.0014$).

B. Images of chlorophyll auto-fluorescing area of *A. thaliana* plants on 5 selected doses of glyphosate.

3.5. Discussion

I have experimentally shown that the liverwort *Marchantia polymorpha* is naturally tolerant to the non-selective herbicide glyphosate, and therefore exhibits one of the only reported examples of non-evolved glyphosate tolerance in plants. In this chapter, I show that this tolerance is not caused by a tolerant form of the EPSPS enzyme in *M. polymorpha*. Therefore, it is likely that there is another mechanism which causes glyphosate tolerance in *M. polymorpha*.

Anecdotal evidence has been reported that *M. polymorpha* is tolerant to glyphosate (Marble et al., 2017). In order to test this experimentally, I conducted a dose-response experiment on *M. polymorpha*, comparing its dose-response relationship to glyphosate with *A. thaliana*. Calculating the IC_{50} values revealed *M. polymorpha* is approximately 12 times more tolerant to glyphosate than *A. thaliana*. Additionally, *M. polymorpha* has a much higher lethal dose of glyphosate compared to *A. thaliana*, with 10 µM glyphosate being lethal to *A. thaliana*, while 100 µM glyphosate is not lethal to *M. polymorpha*. High levels of tolerance may be associated with a TSR mechanism (Gaines et al., 2020).

I conducted a phylogenetic analysis of EPSPS amino acid sequences from land plants to determine if there were differences in MpEPSPS which may cause glyphosate tolerance. The genome of *M. polymorpha* was found to encode a single EPSPS and this formed a well-supported, monophyletic clade with other liverwort EPSPS amino acid sequences included in the analysis. However, liverwort sequences did not form a monophyletic clade with EPSPS sequences from the other bryophytes. Therefore the topology of the EPSPS tree contradicts reported land plant phylogenies, in which bryophytes form a monophyletic clade (Carpenter et al., 2019; Li et al., 2020; Morris et al., 2018). A detailed analysis of the amino acid sequence alignment of land plant EPSPSs revealed no amino acid sequence differences in MpEPSPS in positions known to cause glyphosate resistance and there was a very high level of amino acid conservation across the land plants. The difference between the topologies of reported land plant phylogenies and that of the EPSPS phylogeny I report here, is likely an artifact of this very high amino acid sequence similarity of EPSPS from across the land plants. This is testament to the importance of this enzyme in plants, as any change in amino acid

sequence is likely detrimental to plant health. Structural modelling of EPSPS enzymes from *A. thaliana* and *M. polymorpha* confirmed the high structural conservation in land plants.

Homology models of EPSPS enzymes were used to predict whether, and how, glyphosate binds to MpEPSPS using a molecular docking analysis. Glyphosate was predicted to bind to the EPSPS enzymes from *E. coli*, *A. thaliana* and *Z. mays* in the same conformation, and via at least one covalent bond. These enzymes are glyphosate sensitive. Two EPSPS enzymes used in this analysis, from *Agrobacterium tumefaciens* CP4 strain and the *Zea mays* EPSPS with the resistance conferring TIPS mutation introduced, are glyphosate-resistant and are used in glyphosate resistant crops (Funke et al., 2009; Padgette et al., 1988). Glyphosate was predicated to bind to the homology models of these enzymes in a non-linear, more compact conformation than to the wild-type enzymes and with no covalent bonds. Glyphosate was predicted to bind to the EPSPS from *A. tumefaciens* CP4 in the same conformation that has been reported for X-ray crystal structures of glyphosate-resistant EPSPSs (Funke et al., 2006). This contorted conformation is likely to be less energetically favourable than the natural conformation the molecule usually occurs in, and in which it is predicted to bind to sensitive EPSPS enzymes. Therefore, prediction of binding of glyphosate in this conformation likely represents energetically unfavourable binding which does not occur *in vivo*. Glyphosate is predicted to bind to the homology model of MpEPSPS in the same conformation as is predicted to *A. tumefaciens* CP4 EPSPS and also with no covalent bonds. This suggests that glyphosate binds to MpEPSPS differently than to sensitive EPSPS enzymes, suggesting differences in the MpEPSPS active site. The prediction of equivalent binding as to the glyphosate-resistant EPSPS from *A. tumefaciens* CP4 is consistent with MpEPSPS being glyphosate-tolerant.

My *in silico* analysis predicted MpEPSPS to be glyphosate-tolerant, but subsequent *in vitro* analysis contradicted this prediction. Both crude protein extracts containing MpEPSPS and purified MpEPSPS expressed in *E. coli* exhibited the same dose-response relationship to glyphosate as sensitive EPSPS enzymes from *A. thaliana*, *C. conicum*, and *L. cruciata*. Statistical tests found no significant differences in their IC_{50} values. These results were consistent when compared to the glyphosate-resistant EPSPS enzymes used in glyphosate-resistant crops, which had IC_{50} values approximately 500-fold higher than MpEPSPS.

Therefore, MpEPSPS is not glyphosate-tolerant as initially predicted but rather glyphosate-sensitive.

TSR to glyphosate is usually associated with alterations which reduce the binding affinity of EPSPS for glyphosate. However, it is also formally possible that differences in the kinetic properties of an EPSPS enzyme impact glyphosate tolerance at the whole-plant level. Kinetic analysis of the purified EPSPS enzymes analysed in the dose-response assay revealed the most significant differences between bacterial and plant EPSPS enzymes, regardless of their glyphosate sensitivity or resistance. Furthermore, differences in kinetic properties were also observed among wild-type plant EPSPS enzymes. These differences may be sufficient to influence glyphosate tolerance at the whole-organism level.

The bacterial EPSPS enzymes tested, from *E. coli* and *A. tumefaciens* CP4 were found to have far higher affinities for PEP than the plant EPSPS enzymes. This implies the ability of these enzymes to efficiently utilise PEP as a substrate. Since glyphosate competes for the binding site of PEP, the high affinity of the bacterial EPSPS enzymes for PEP also implies the inherent higher sensitivity of these enzymes to glyphosate (Gaines et al., 2011). However, the *A. tumefaciens* CP4 EPSPS is highly glyphosate resistant, suggesting the amino acid differences causing this resistance do not greatly diminish the enzyme's affinity to PEP, but do so specifically to glyphosate. Wild-type plant EPSPSs were found to have a lower affinity for PEP, and all the plant EPSPSs were far more sensitive to glyphosate than the EPSPS from *A. tumefaciens* CP4. Interestingly, the introduction of the double TIPS mutation into the *Z. mays* EPSPS did not significantly alter its affinity for PEP, suggesting that this mutation does not directly affect the ability of the EPSPS to bind PEP. However, this mutation does have a severe impact on the maximum velocity (V_{max}) and catalytic efficiency (K_{cat}/K_m) of the enzyme. This indicates that the TIPS mutation may not inhibit the enzyme's ability to bind PEP but may affect other aspects of the enzyme's catalytic activity, leading to its reduced efficiency of utilising PEP as a substrate. However, this catalytic activity must be sufficient to allow plant survival in glyphosate tolerant crops utilising this enzyme in presence of glyphosate.

The affinities of the wild-type plant EPSPS enzymes for PEP and S3P are similar, and therefore this likely does not contribute to the natural glyphosate tolerance in *M.*

polymorpha. However, the V_{max} values for each substrate do differ between the enzymes. The MpEPSPS enzyme has the lowest V_{max} when varying both PEP and S3P. This lower V_{max} may contribute to *M. polymorpha's* natural tolerance to glyphosate at the whole plant level. *M. polymorpha* may have evolved an EPSPS with a slower rate of catalysis in order to use the shikimate pathway more conservatively compared to other plant species with higher V_{max} values for EPSPS. This may be advantageous in resource limited environments, where the plant allocates energy and resources more efficiently, contributing to its survival in such habitats. This may allow *M. polymorpha* to survive glyphosate treatment, which essentially limits resources for the plant, since approximately 20 % of photosynthetically fixed carbon flows through the shikimate pathway in plants (Maeda and Dudareva, 2012). The lower V_{max} of MpEPSPS likely reflects an evolutionary trade-off, balancing the need for shikimate pathway activity, while minimising the cost of enzyme production and turnover. It is then also likely that the evolved decrease in V_{max} has been accompanied by changes in the regulation of the shikimate pathway in *M. polymorpha* or in the plant's utilisation of alternate sources of aromatic amino acids during stress situations. If lower catalytic efficiency in MpEPSPS contributes to glyphosate tolerance, it could have significant implications for weed management. Weeds with EPSPS enzymes exhibiting lower catalytic activity might survive higher glyphosate doses, potentially providing a selective advantage and facilitating resistance evolution. To test this, the inherent glyphosate tolerance of weed plants could be measured and compared to the catalytic activity of their EPSPS enzymes.

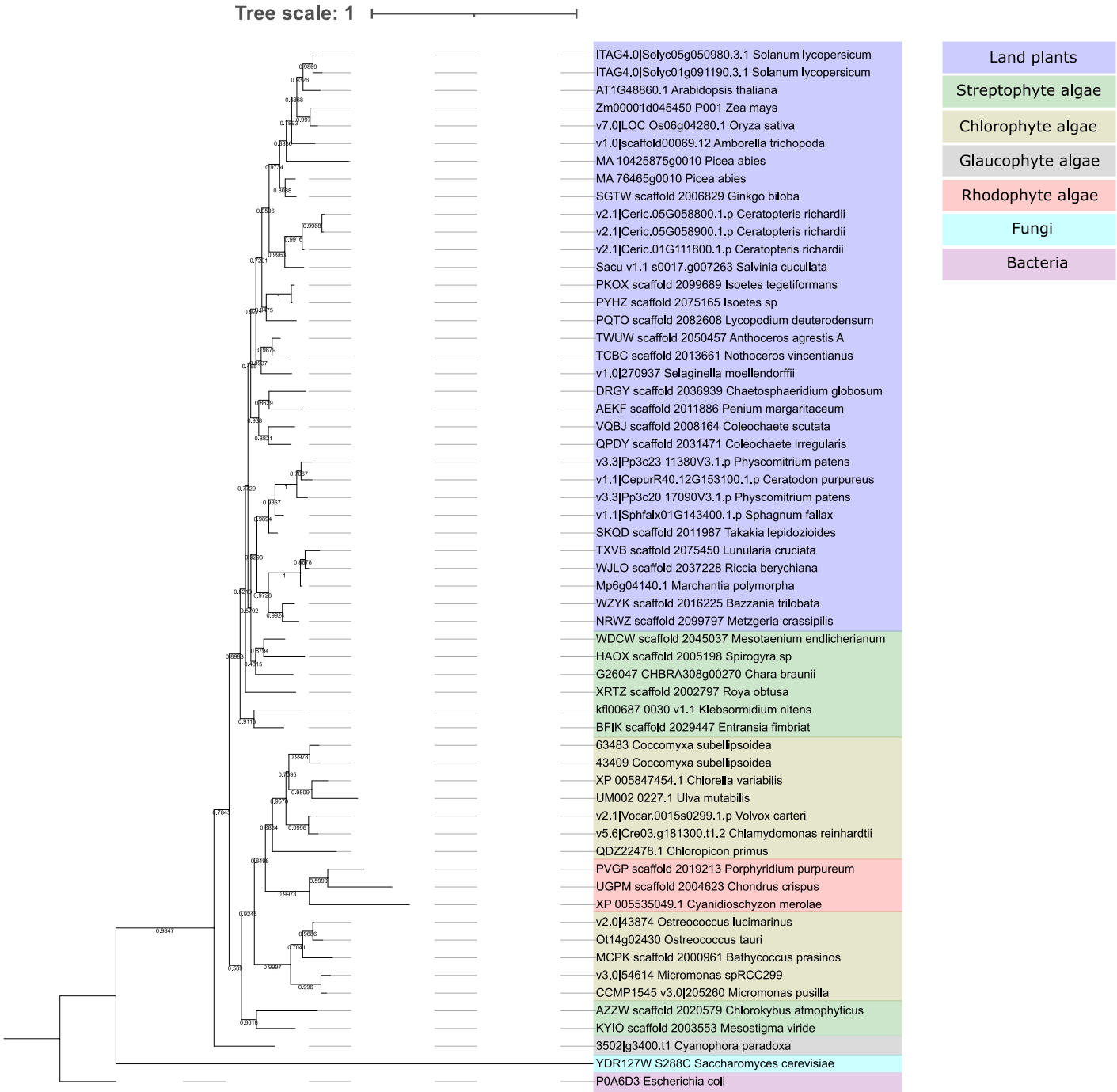
When overexpressed in the glyphosate sensitive *A. thaliana*, MpEPSPS does not increase glyphosate resistance more than overexpression of the endogenous EPSPS. This likely reflects the lower catalytic activity of the *M. polymorpha* EPSPS enzyme. This evidence does not however, refute the hypothesis that the low activity of MpEPSPS may contribute to glyphosate tolerance in *M. polymorpha* by the evolution of compensatory traits.

I identified a serine at position 500 in MpEPSPS which may lead to the reduced activity of the enzyme. This residue is directly linked to 3 residues which form the active site of EPSPS. By changing the residue at this position in MpEPSPS to a valine (the residue present in this position in *A. thaliana*), the resulting homology model of this amino acid sequence was then found to bind glyphosate, in the same, linear conformation as it is predicted to bind to other land plant EPSPSs. This evidence is consistent with this residue causing alterations to the

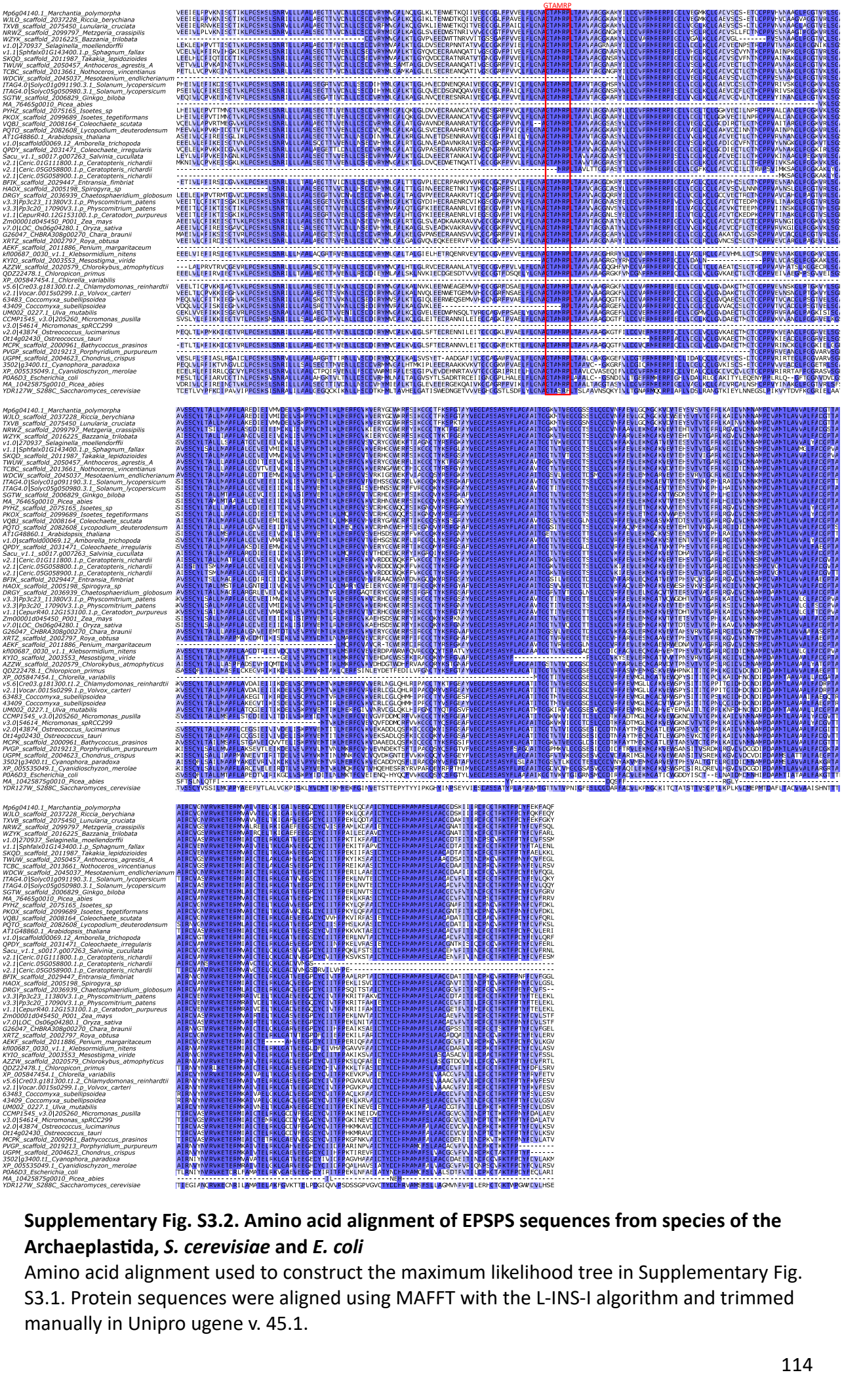
active site of MpEPSPS, which may lead to its lower activity – potentially contributing to glyphosate tolerance in *M. polymorpha*. Overexpression in *A. thaliana* of altered forms of AtEPSPS and MpEPSPS, with the residue at position 493 and 500 exchanged respectively, led to a reduction in the level of glyphosate resistance in both cases. This suggests any amino acid change at this position reduces EPSPS activity, supporting its role in active site formation and catalytic activity. However, this evidence does not support that the serine in MpEPSPS is solely responsible for its lower activity.

To summarise, *M. polymorpha* exhibits high levels of glyphosate tolerance, however this is not due to a tolerant EPSPS enzyme. *M. polymorpha* therefore does not likely employ a known mechanism of TSR, in which glyphosate does not bind efficiently to EPSPS. In my next chapter, I identify a novel tolerance mechanism which may account for the majority of the glyphosate tolerance in *M. polymorpha*.

3.6. Supplementary data

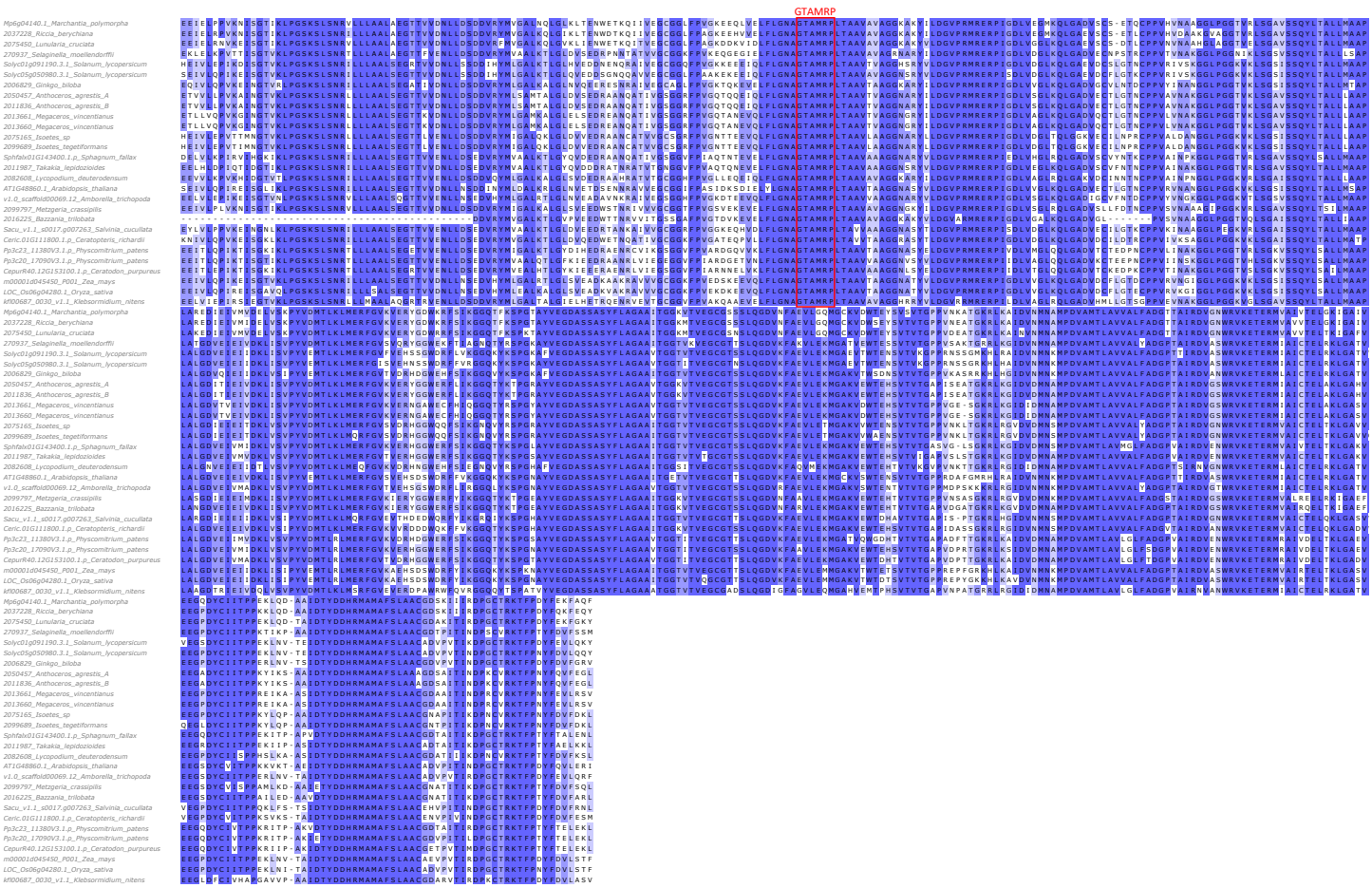


Supplementary Fig. S3.1. EPSPS is very highly conserved throughout the Archaeplastida
Maximum likelihood tree of EPSPS amino acid sequences from species from across the Archaeplastida, *S. cerevisiae* and *E. coli*, rooted to the EPSPS sequence from *E. coli*. Phylogenetic tree was constructed using PHyML 3.0 (Guindon et al., 2010) using an estimated gamma distribution model and the WAG+G+I model of amino acid substitution. Support values are given from a Chi^2-based aLRT test. Protein sequences were aligned using MAFFT with the L-INS-I algorithm.



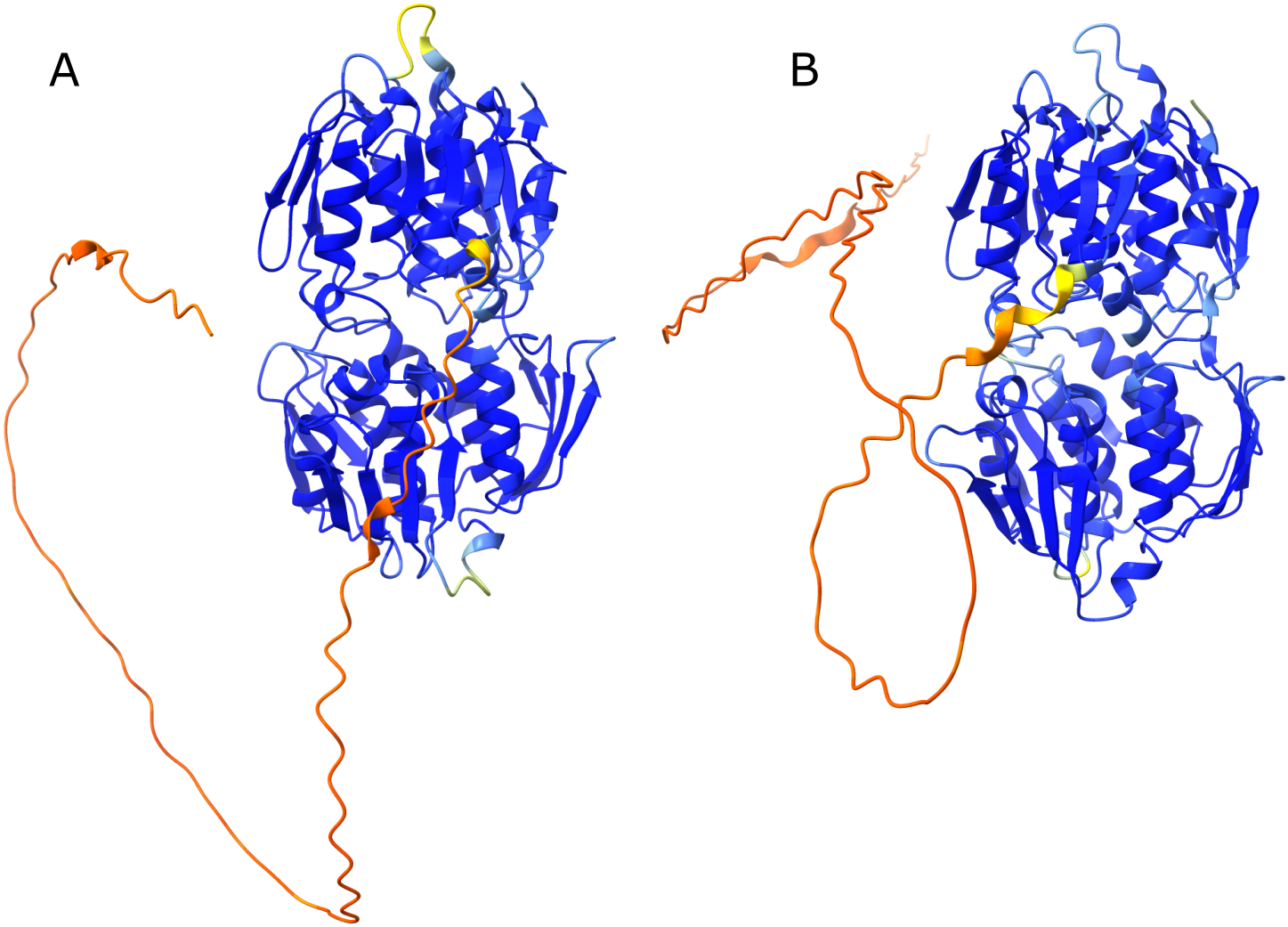
Supplementary Fig. S3.2. Amino acid alignment of EPSPS sequences from species of the Archaeplastida, *S. cerevisiae* and *E. coli*

Amino acid alignment used to construct the maximum likelihood tree in Supplementary Fig. S3.1. Protein sequences were aligned using MAFFT with the L-INS-I algorithm and trimmed manually in Unipro ugene v. 45.1.



Supplementary Fig. S3.3 Amino acid alignment of EPSPS sequences from species of the land plants

Amino acid alignment used to construct the maximum likelihood tree in Fig. 3.3. Protein sequences were aligned using MAFFT with the L-INS-I algorithm and trimmed manually in Unipro ugene v. 45.1.



Supplementary Fig. S3.4. AlphaFold prediction of *A. thaliana* and *M. polymorpha* EPSPS enzymes

A. *A. thaliana* EPSPS.

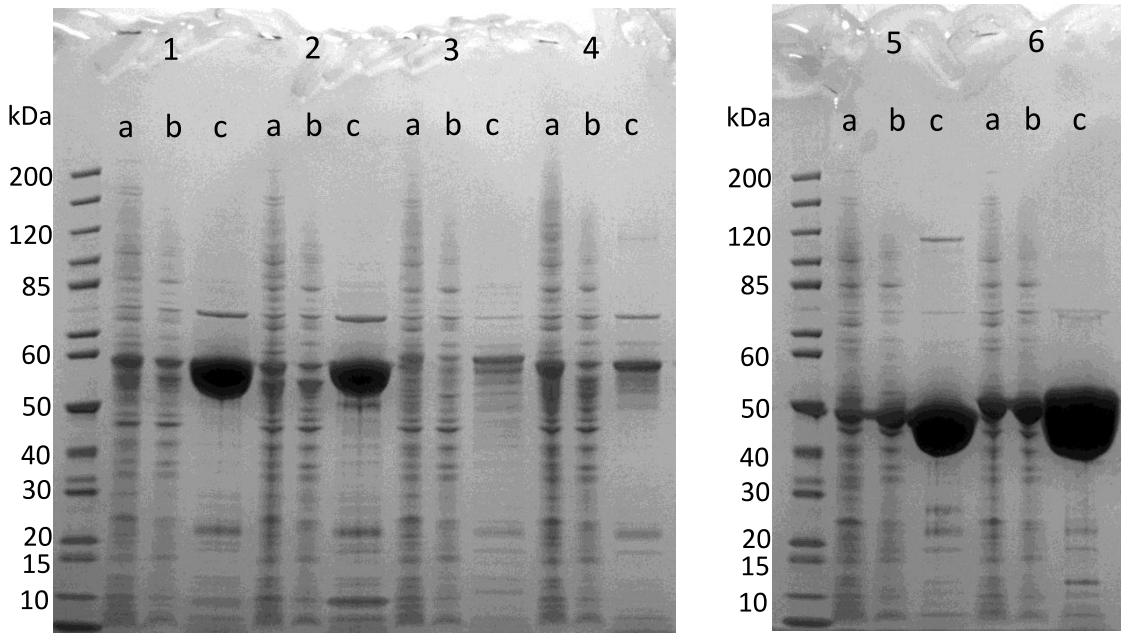
B. *M. polymorpha* EPSPS



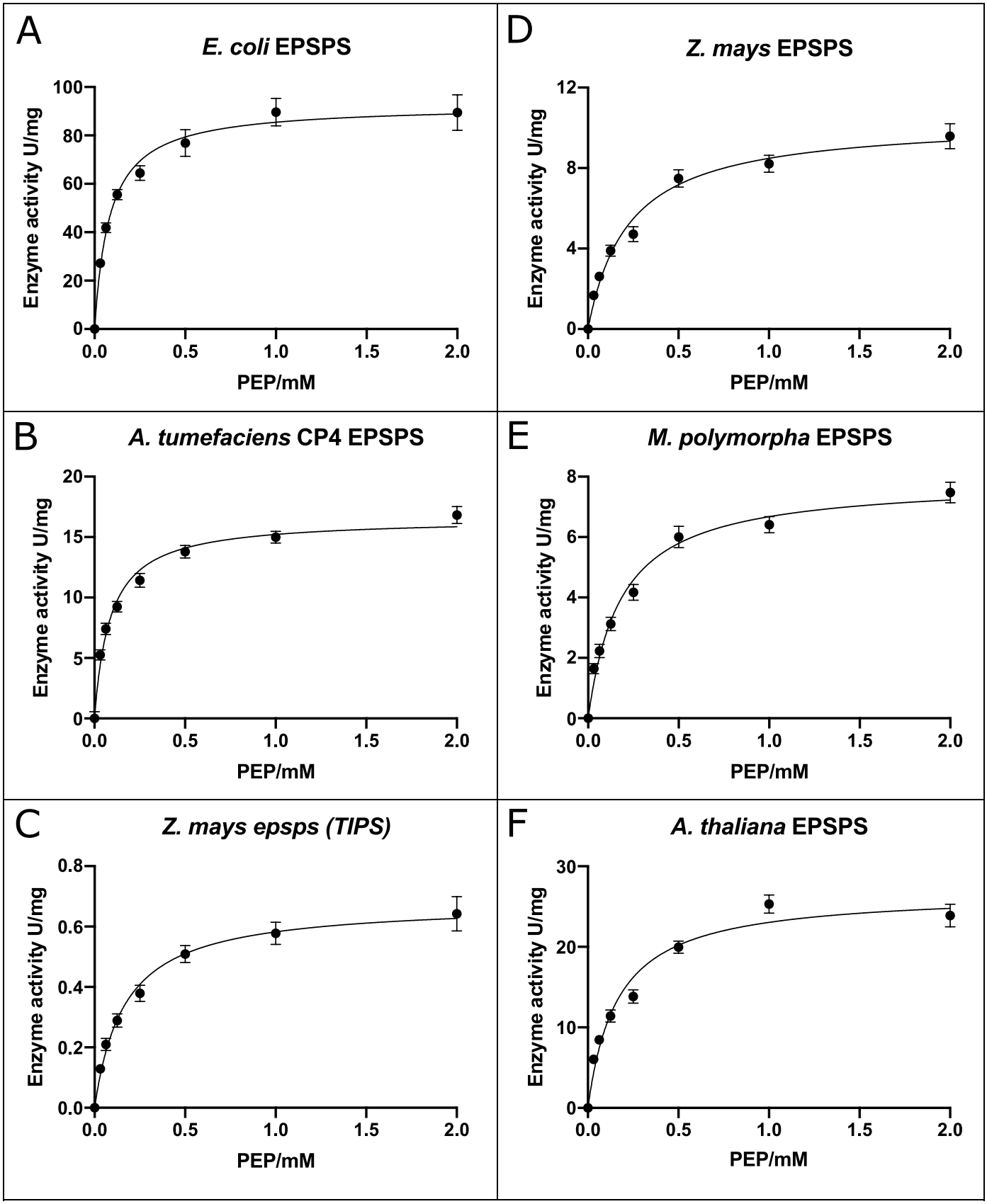
Supplementary Fig. S3.5. Liverwort species from which crude protein extracts were taken
A. *Marchantia polymorpha*. B. *Conocephalum conicum* C. *Lunularia cruciata*

Number	Name	MW (kDa)
1	MAIZE_EPSPS	57.2
2	MAIZETIPS_EPSPS	57.5
3	MARPO_EPSPS	59.9
4	AROA_ARATH_EPSPS	59.6
5	AROA_ECOLI_EPSPS	49.9
6	AROAAGRSC_CP4_EPSPS	51.4

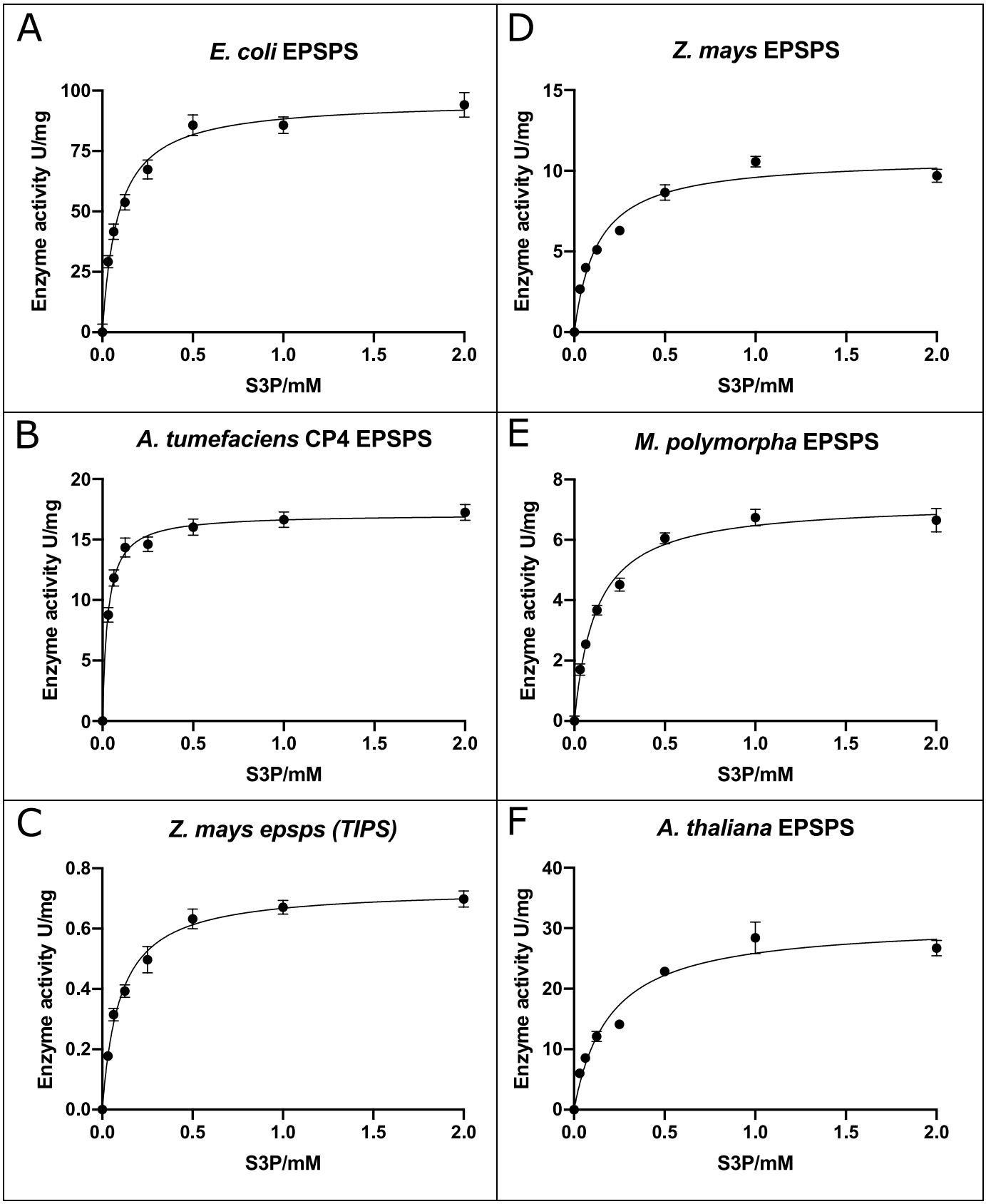
a. WC- whole cell lysate
b. SN- supernatant
c. Elution



Supplementary Fig. S3.6. Expression and purification of EPSPS enzymes



Supplementary Fig. S3.7. Michaelis-Menten curves of EPSPS enzymes to determine kinetic properties in relation to PEP



Supplementary Fig. S3.8. Michaelis-Menten curves of EPSPS enzymes to determine kinetic properties in relation to S3P

Chapter 4: The peptidoglycan biosynthesis enzyme UDP-N-acetylglucosamine-3-enopyruvyl transferase (MurA) confers glyphosate tolerance in *Marchantia polymorpha*

4.1 Abstract

Glyphosate is a broad-spectrum, non-selective herbicide, and therefore natural tolerance to glyphosate is rare among angiosperms. However, the liverwort *Marchantia polymorpha* can survive concentrations of glyphosate many times that which is lethal to angiosperms. The target of glyphosate, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), is sensitive to glyphosate in *M. polymorpha*. This suggests that a known target-site resistance (TSR) mechanism does not confer glyphosate tolerance in *M. polymorpha*. I identify the EPSPS-like enzyme UDP-N-acetylglucosamine-3-enolpyruvyl transferase (MurA) as a candidate contributing to glyphosate tolerance in *M. polymorpha*, because it is structurally similar to EPSPS. I show that MurA is absent from glyphosate-sensitive angiosperms but well conserved among non-angiosperm streptophytes, including bryophytes. Consistent with a role of MurA in glyphosate tolerance, I show that two bryophytes which encode MurA, *Physcomitrium patens* and *Lunularia cruciata*, are also tolerant to glyphosate. Furthermore, overexpression of MurA in *M. polymorpha* increases glyphosate resistance, while *murA* loss-of-function mutant plants are more sensitive to glyphosate. I found that *murA* loss-of-function mutant plants exhibit a macrochloroplast phenotype, due to a disruption of peptidoglycan biosynthesis leading to chloroplast division defects. To determine if MurA confers glyphosate tolerance independently of peptidoglycan biosynthesis, I generated loss-of-function mutants of another peptidoglycan biosynthesis enzyme, MurF, and treated wild-type plants with peptidoglycan biosynthesis inhibiting antibiotics. A macrochloroplast phenotype was observed in both instances, but there was no change in sensitivity to glyphosate, suggesting that peptidoglycan does not play a role in glyphosate tolerance in *M. polymorpha*. Consistent with this, *A. thaliana* plants expressing MpMurA were more than 7 times more resistant to glyphosate than control plants. I then show that MpMurA catalyses the same reaction as EPSPS, transferring the enolpyruvyl moiety of PEP to shikimate-3-phosphate forming EPSP product. I propose that MurA may therefore confer glyphosate tolerance either through a mechanism analogous to target-site overexpression or by binding glyphosate, making it less available to bind to EPSPS. This represents a novel mechanism of herbicide tolerance.

4.2. Introduction

Marchantia polymorpha is tolerant to glyphosate. This is not caused by a glyphosate-tolerant variant of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), the glyphosate target. The glyphosate sensitivity of EPSPS in *M. polymorpha* is consistent with glyphosate tolerance being the result of a form of non-target site resistance (NTSR). In this chapter, I show that the enzyme UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) contributes to glyphosate tolerance in *M. polymorpha*.

Evolved herbicide resistance in weeds can be broadly classified into two groups: target site resistance (TSR) and non-target site resistance (NTSR) (Gaines et al., 2020). TSR occurs due to mutations in the gene encoding the target of the herbicide, or those which lead to overexpression of the target protein (Ghanizadeh et al., 2019; Murphy and Tranel, 2019). Overexpression of a herbicide target can confer herbicide resistance by increasing the concentration of the target protein within the plant. This reduces the proportion of targets bound with the herbicide, subsequently reducing its inhibitory effects on the plant (Gaines et al., 2020). Target overexpression has evolved as a mechanism of resistance to several herbicide classes. For example, overexpression of Acetyl-CoA carboxylase (ACCase) was found to confer resistance to ACCase inhibiting herbicides such as fenoxaprop-P-ethyl, in resistant populations of large crabgrass (*Digitaria sanguinalis*) (Laforest et al., 2017). Similarly, overexpression of the gene encoding acetolactate synthase (ALS) has evolved in populations of *Bromus sterilis* to give resistance to ALS-inhibiting herbicides (Sen et al., 2021). Overexpression of the glyphosate target EPSPS has also been reported as one of the most prevalent mechanisms of resistance against this herbicide (Chen et al., 2020; Gaines et al., 2010; Molin et al., 2020; Salas et al., 2012; Yanniccari et al., 2023). Overexpression of EPSPS confers glyphosate resistance as there is a sufficient proportion of unbound EPSPS enzymes to allow carbon flux through the shikimate pathway and for aromatic amino acids to be ultimately synthesised. It is also formally possible that the expression of a structurally or functionally similar protein to the herbicide target – rather than the target itself – could have a similar effect. Such a mechanism would be classified as NTSR.

NTSR mechanisms reduce the amount of herbicide reaching the target or protect the plant from the toxic effects of the herbicide downstream of its action (Jugulam and Shyam, 2019). The most studied genes which contribute to NTSR in weeds encode proteins from four enzyme families: cytochrome P450s, glutathione S-transferases, glycosyltransferases, and ATP-binding cassette transporters (Yu and Powles, 2014; Yuan et al., 2007). However, there are various other NTSR mechanisms plants can evolve to herbicides. For example, decreased herbicide absorption and/or translocation, sequestration of the herbicide to where it is isolated from its target, and mechanisms which bypass the inhibition of the target or its pathway (Gaines et al., 2020). I hypothesise that an NTSR mechanism is responsible for glyphosate tolerance in *Marchantia polymorpha*.

Glyphosate is a foliar applied, systemic herbicide, which is translocated from the leaves to the growing tissues of the plant where it inhibits its target, EPSPS, in the chloroplast. Many populations of weed species have evolved glyphosate resistance through mechanisms of reduced translocation (Sammons and Gaines, 2014). For example, in populations of *Lolium rigidum* (Lorraine-Colwill et al., 2002)*, Conyza canadensis* (Feng et al., 2004), and *Ambrosia trifida* (Moretti et al., 2018; Van Horn et al., 2018). Resistance may also evolve through mechanisms which reduce the initial absorption of glyphosate (Gaines et al., 2020). Such mechanisms have been proposed to be attributed to differences in plant cuticle thickness and/or composition (Hess, 1985). Glyphosate is primarily used to control angiosperm weeds and is generally effective towards these species. Therefore, the glyphosate tolerance observed in *M. polymorpha* may be a result of fundamental differences in morphology and vasculature between liverworts and angiosperms, which may not facilitate efficient translocation of glyphosate to its target in *M. polymorpha*.

One difference between *M. polymorpha* and angiosperm weeds which may affect glyphosate translocation is the predicted presence of a peptidoglycan wall surrounding chloroplasts in *M. polymorpha* (Machida et al., 2006; Tounou et al., 2002). The *M. polymorpha* genome contains nine genes essential for peptidoglycan biosynthesis, whereas the model angiosperm *A. thaliana* only contains five genes and there is limited evidence for the formation of a peptidoglycan wall in vascular plants (Machida et al., 2006). Peptidoglycan surrounds the chloroplast in some green plants such as *Physcomitrium*

patens, Closterium algae and *Klebsormidium nitens* (Dowson et al., 2022; Hofmann, 2016; Homi et al., 2009; Matsumoto et al., 2012; Sato et al., 2017; Takano et al., 2018). The genes encoding peptidoglycan biosynthesis enzymes are also required for chloroplast division, with loss-of-function mutations in these genes leading to chloroplast division defects (Katayama et al., 2003; Machida et al., 2006; Utsunomiya et al., 2021). Given that glyphosate targets EPSPS, which is chloroplast localised, differences in chloroplast wall composition and chloroplast division mediators between *M. polymorpha* and angiosperms may contribute to glyphosate tolerance in *M. polymorpha*.

In this chapter, I aim to identify the mechanism which confers glyphosate tolerance in *M. polymorpha*. I predict that this tolerance mechanism will not involve MpEPSPS. Using a BLASTP search, I identify the EPSPS-like enzyme UDP-N-acetylglucosamine-3-enopyruvyl transferase (MurA) in the proteome of *M. polymorpha*. If MurA contributes to glyphosate tolerance in *M. polymorpha*, I predict that MurA is not present in angiosperms, which are generally sensitive to glyphosate. Using a BLASTP search of the proteomes of Archaeplastida species, I show MurA to be highly conserved among non-angiosperm streptophytes but absent from angiosperms. Endogenous overexpression of MurA in *M. polymorpha* causes an increase in glyphosate resistance while Mp*murA* loss-of-function mutants, generated using the CRISPR-Cas9 mutagenesis system, are more sensitive to glyphosate compared to control plants. This evidence is consistent with MurA contributing to glyphosate tolerance in *M. polymorpha*. The contribution of MurA to glyphosate tolerance was independent of peptidoglycan biosynthesis as Mp*murF* loss-of-function mutants were not more sensitive to glyphosate and exogenous overexpression of MpMurA in *A. thaliana* confers glyphosate resistance in this angiosperm. MurF is another enzyme involved in peptidoglycan biosynthesis. I then show that MpMurA catalyses the conversion of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) into 5-enolypyruvylshikimate-3-phosphate (EPSP) and phosphate, like MpEPSPS. The activity of MpMurA with the substrates S3P and PEP may reduce the inhibitory effects of glyphosate treatment in *M. polymorpha* and represents a novel mechanism of glyphosate tolerance.

4.3. Materials and Methods

4.3.1. Protein Structure Prediction using AlphaFold

EPSPS and MurA protein structures from M. polymorpha were predicted using AlphaFold (Jumper et al., 2021). Sequences of EPSPS (Mp6g04140.1) and MurA (**Mp5g14110.1) were retrieved from** MarpolBase (http://marchantia.info/) and used as inputs for AlphaFold.

To compare these structures, ChimeraX (Pettersen et al., 2021) was employed. EPSPS and MurA structures were loaded as separate PDB files, and the MatchMaker tool used to align them, minimising the root-mean-square deviation (RMSD) for structural similarity assessment.

4.3.2. Phylogenetic analysis of *MurA* genes

See Chapter 3 Section 3.1.

4.3.3. *M. polymorpha* spore sterilisation

Freshly harvested spores were sterilised with 1 % NaDCC (sodium dichloroisocyanurate) for 3 to 5 minutes. Spores were then washed with sterile dH_2O.

4.3.4. Plant lines and growth conditions

See Chapter 3 Section 3.4.

4.3.5. Construct making

DNA sequences of full and part length *M. polymorpha MurA* genes were synthesised by Twist Biosciences. Synthesised products were then cloned using GreenGate cloning (Lampropoulos et al., 2013). Products were cloned into entry module pGGI000, transformed into *E. coli*, and purified using the Vienna Biocenter's in-house kit. Subsequently, this plasmid, the promotor module (35S), and the terminator module (RBCS) were combined with the destination vector (pGGZ003). Resulting vectors were introduced into *A. tumefaciens* through electroporation.

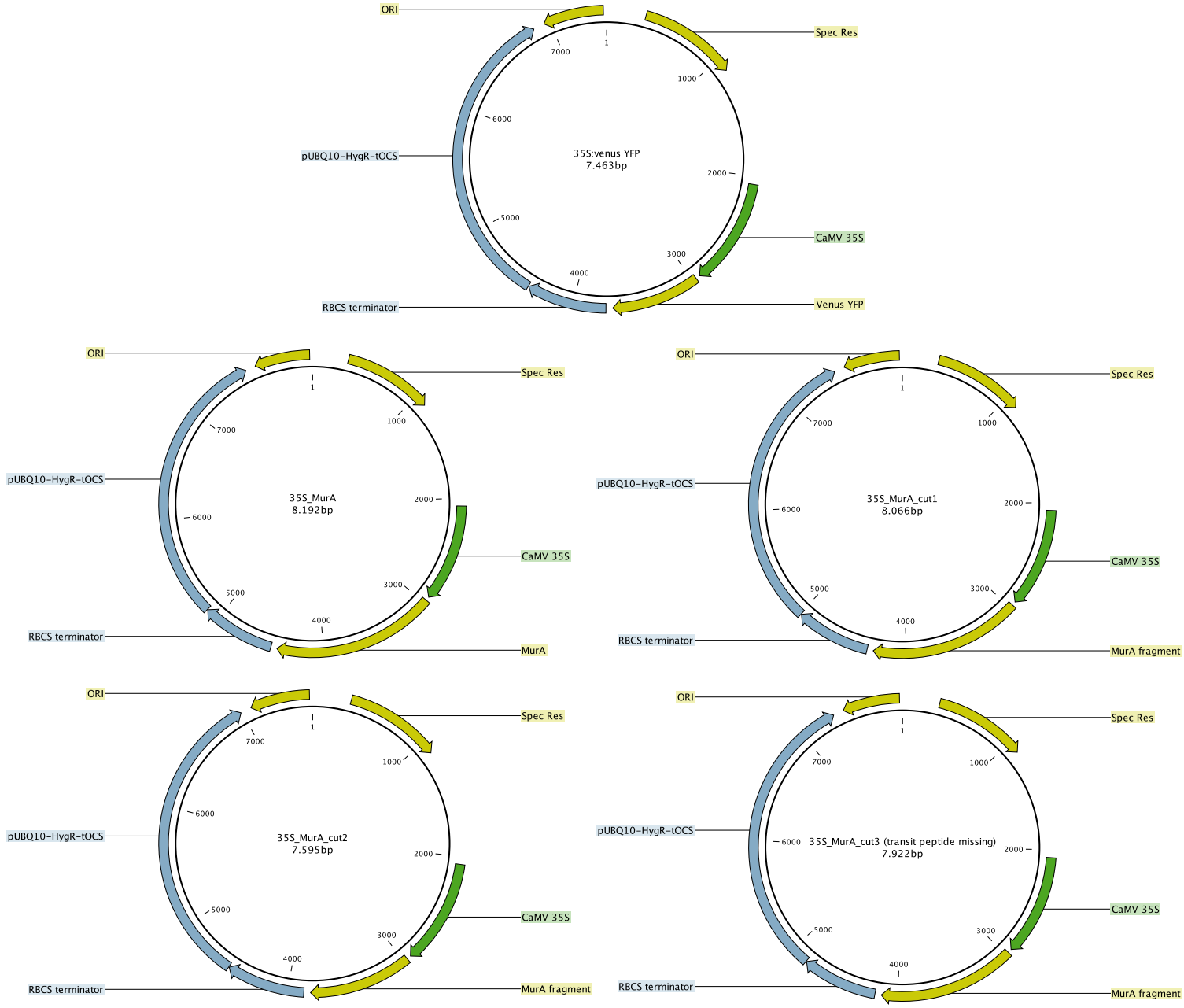


Fig. 4.1. Vector map of *MurA* gene expression constructs
Full and part length DNA sequences of the *MurA* gene from *M. polymorpha* were cloned into a vector using GreenGate cloning. The same vector with the YFP protein in place of *MurA* was used as a control

4.3.6. Guide RNA design

M. polymorpha MurA (**Mp5g14110.1)** and *MurF* (Mp1g21490) genes were mutagenised using CRISPR-Cas9. Guide RNAs targeting specific gene regions with "NGG" sites (PAM sequence) were designed in the 5' direction (Sugano et al., 2014). The following guide RNA sequences were used (PAM sequences in parentheses):

murA sgRNA 1	AGCAAACTTTCTGGCCATGT(CGG)
murA sgRNA 2	TTTCGTACAGGCCAGCAGG(AGG)
murF sgRNA 1	GATAGTCGGAAAATCAACCG(AGG)
murF sgRNA 2	ATAGTCGGAAAATCAACCGA(GGG)

4.3.7. Cloning of CRISPR-Cas9 vectors

CRISPR-Cas9 vectors were cloned using the OpenPlant toolkit which utilises Golden Gate cloning methods (Sauret-Güeto et al., 2020). Oligos were annealed using the following reaction components and conditions:

1 µl gRNA oligo F (100 µM)

1 µl gRNA oligo R (100 µM)

8 µl ddH_2O

The oligos were annealed in a thermo-cycler for 30 minutes at 37 ℃, followed by 5 minutes at 95 ℃ then ramp down to 25 ℃ at 5 ℃ per minute.

The vector OP-076 was assembled in the following reaction:

ddH_20	5µl
10x Tango buffer (Thermo Fisher)	1µl
1mg/ml bovine serum albumin (NEB)	0.5µl
T4 DNA ligase (5U/µl, NEB)	0.25µl
*Sap*I (*Lgu*I) (5U/µl, Thermo Fisher)	0.25µl
10mM ATP (SIGMA)	1µl
OP-76 (25-50ng)	1µl
annealed oligo	1µl
total	10µl

The reaction was run in a thermo-cycler with 15 cycles of 3 minutes at 37 ℃ and 4 minutes at 16 ℃, then 5 minutes at 50 ℃ and 10 minutes at 80 ℃.

Resulting vectors were transformed into DH5-α *E. coli* cells by electroporation and plated on LB media with 50 µg ml^{-1} spectinomycin and 40 µg ml^{-1} X-gal. After an overnight incubation at 37 ℃, positive colonies were selected, cultured in LB media containing 50 µg ml^{-1} spectinomycin, and plasmids were extracted using a MiniPex 3 in 1 kit. Sanger sequencing with the Mp-U6-seqF primer (Section 4.3.17) confirmed proper assembly of the gRNA oligo duplex into the CRISPR expression vector.

4.3.8. *Agrobacterium tumefaciens*-mediated transformation of *M. polymorpha* spores

Vectors were introduced into *M. polymorpha* spores using a modified method from Ishizaki et al., 2008. *Agrobacterium* strain GV3101 was transformed with vectors via electroporation. This strain is resistant to rifampicin and gentamycin, and successful transformants exhibited hygromycin resistance. Transformed *A. tumefaciens* were cultured on solid LB media with 50 µg ml^{-1} hygromycin, gentamycin, and rifampicin for 2 days at 28 ℃ to select colonies incorporating the overexpression vectors. Single colonies were cultivated in 5 ml liquid LB media supplemented with hygromycin, gentamycin, and rifampicin, then induced with 100 µM acetosyringone in M51C media for 3-6 hours at 28 ℃, shaking at 150 rpm. 1 ml of the induced culture was combined with a 25 ml culture of *M. polymorpha* spores (resulting from a cross between Tak-1 and Tak-2 accessions) grown for 5 days in M51C media at 23 ℃ under 10-30 µmol m^{-2} s^{-1} white light. The spore and *Agrobacterium* culture was co-cultivated for 2 days at 23 ℃, shaking at 120 rpm, with 100 µM acetosyringone. Then, the culture was filtered through a 50 µM cell strainer to collect the sporelings, with *Agrobacterium* removed by washing with 150 ml dH2O. Sporelings were plated on solid ½ Gamborg medium with 1% agar, supplemented with 100 µg ml^{-1} cefotaxime (to eliminate any remaining *Agrobacteria*) and 10 µg ml^{-1} hygromycin for selection. The selection occurred at 23 ℃ under 10-30 µmol m^{-2} s^{-1} white light for 1-2 weeks. Surviving plants were then transferred to new plates with solid ½ Gamborg medium with 1% agar, supplemented with 100 µg ml^{-1} cefotaxime and 10 µg ml^{-1} hygromycin, for an additional week.

4.3.9. Agrobacterium tumefaciens-mediated transformation of *A. thaliana* plants

See Chapter 3 Section 3.6.

4.3.10. Genotyping putative mutants generated by CRISPR/Cas9 mutagenesis

Putative Mp*murA* and Mp*murF* mutants were genotyped using the Phire Plant Direct PCR kit (Thermo Scientific™). Primers were designed to amplify gene fragments containing expected mutations by binding to the gene of interest flanking the sgRNA binding sites (Section 4.3.17). The resulting PCR product underwent purification using ExoSAP treatment to eliminate primers and dNTPs. The ExoSAP mix included 0.04 µl Exonuclease I (NEB Cat. M0293), 0.4 µl Shrimp Alkaline Phosphotase (NEB Cat. M0371), and 1.56 µl of storage solution (Section 4.3.18) for every 20 µl of PCR product. Purified PCR products were sequenced using Sanger sequencing and aligned with reference Mp*MurA* and Mp*MurF* sequences in CLC Genomics Workbench 9 to identify mutations introduced.

4.3.11. Phenotyping putative mutants generated by CRISPR/Cas9 mutagenesis and plants treated with peptidoglycan biosynthesis inhibiting antibiotics

Gemmae were grown for 3 days on solid ½ Gamborg media (Section 4.3.18) with or without relevant antibiotics. Antibiotic doses were the maximum tolerable doses without severe symptoms after 14-day treatment. Gemmalings were imaged using a Keyence VHX-7000 microscope, with Tak-1 gemmae grown on non-supplemented media serving as controls. FIJI was used to open and quantify chloroplast abundance and chloroplast dimensions in the images (Schindelin et al., 2012). Statistical analysis was conducted using GraphPad Prism 9.3.1.

4.3.12. Herbicide and stock solution preparation

See Chapter 3 Section 3.7.

4.3.13. Dose-response assays

See Chapter 3 Section 3.8.

4.3.14. MurA and EPSPS expression and purification

MurA and EPSPS proteins from *M. polymorpha* were expressed and purified following the methods outlined in Chapter 3, Section 3.11. MpMurA had a molecular weight of 50 kDa.

4.3.15. MurA and EPSPS enzyme assays

Activities of purified MpEPSPS and MpMurA enzymes were measured using the same method as in Chapter 3, Section 4.5. The reaction consisted of 1 µM MurA enzyme and 3.2 nM EPSPS enzyme in an 80 µl assay.

The amount of EPSP produced in reactions involving MurA or EPSPS with S3P and PEP substrates was quantified using hydrophilic interaction liquid chromatography (HILIC) coupled to liquid chromatography-tandem mass spectrometry (LC-MS/MS) at the Vienna BioCenter Core Facilities Metabolomics Facility. Each reaction (1 µl) was injected onto a polymeric iHILIC-(P) Classic HPLC column (HILICON, 200 x 2.1 mm; 5 µm) with a guard column, operated at a flow rate of 100 µl/min. Separation utilised a linear gradient (A: acetonitrile; B: 10 mM aqueous ammonium bicarbonate with 0.1 µg/ml medronic acid) starting at 25% B and reaching 90% B in 14 minutes. Mass spectrometry was performed using a TSQ Quantiva mass spectrometer (Thermo Fisher Scientific).

EPSP was identified by characterising the product of the in vitro reaction: phosphoenolpyruvate (PEP) + shikimate-3-phosphate (S3P) ⇌ phosphate + 5-enolpyruvylshikimate-3-phosphate (EPSP). A compound with an m/z value of 323.1 was detected in negative ion mode, matching the calculated m/z value of EPSP. Fragmentation confirmed its identity, and this compound was absent in negative controls. Glyphosate sensitivity of MpMurA and MpEPSPS was tested with 1 mM S3P and PEP along with varying glyphosate doses.

4.3.16. Micro-scale thermophoresis binding assays

Micro-scale thermophoresis (MST) measurements were carried out using a NanoTemper Monolith machine. Purified *M. polymorpha* EPSPS and MurA proteins were labelled using the NanoTemper RED-tris-NTA 2nd generation HIS- labelling kit (subsequently called ‘dye’), following the kit instructions (2:1 labelling of the proteins). 200 nM protein was mixed with

100 nM dye and incubated for 30 mins in the dark at RT according to the labelling kit instructions. Labelled protein was measured and compared to dye alone to confirm labelling. PBS buffer with 0.05 % Tween-20 added (PBS-T) was used for dilutions and measurements. The proteins were filtered using 0.45 µM Ultrafree®-MC-HV Durapore® PVDF spin filters. Protein concentration was determined by measuring the scattering corrected 280 nm absorbance in a 0.15 cm cuvette and assuming an extinction coefficient of ε_{280nm} = 38700 $M^{-1}cm^{-1}$ or ε_{280nm} = 3041 $M^{-1}cm^{-1}$. Ligands added in a 16 point 1:1 dilution series with 1 mM as the highest concentration, with or without the presence of 1 mM of S3P. All experiments with EPSPS were carried out with an MST power of 20 %, while those with MurA were carried out with a power 40 %. MST traces were evaluated at a hot time between 0.5 sec and 1.5 sec after turning on the IR laser. Premium capillaries were used throughout the experiment. Resulting binding curves were analysed with PALMIST (Thomas H. Scheuermann et al., 2016; Tso et al., 2018).

4.3.17. Primers

Primer	Sequence	Use
Mp-U6-seqF	GAGTTCGGATTGCTCTCTTTC	Sanger sequencing of CRISPR expression vector
MurA1_fwd	ACCGAGCTCTTGGATCTCAGC	Genotyping Mp*murA* lines (sgRNA1, sgRNA2, sgRNA3); sequencing Mp*murA* lines (sgRNA1, sgRNA2, sgRNA3)
MurA1_rev	ACGCAGAGTTAGCAGCGCTCG	Genotyping Mp*murA* lines (sgRNA1, sgRNA2, sgRNA3)
MurA2_fwd	AGAGAGCTGGACCTTTATGC	Genotyping Mp*murA* lines (sgRNA4); sequencing Mp*murA* lines (sgRNA4)
MurA2_rev	TGACATAGAAATGGAGGACCG	Genotyping Mp*murA* lines (sgRNA4)
MurA3_fwd	CAATTCACTTGTCTGGATGC	Genotyping Mp*murA* lines (sgRNA5); sequencing Mp*murA* lines (sgRNA5)
MurA3_rev	TCAAATGATTCATAACCTCG	Genotyping Mp*murA* lines (sgRNA5)
MurF1_fwd	TCTACCGATCGTTCACATGC	Genotyping Mp*murF* lines (sgRNA2, sgRNA3); sequencing Mp*murF* lines (sgRNA2, sgRNA3)
MurF1_rev	CTCAAAGTTCTCCATATGCG	Genotyping Mp*murF* lines (sgRNA2, sgRNA3)
MurF2_fwd	TCATGTCAAGCTTGTGGAAGC	Genotyping Mp*murF* lines (sgRNA1); sequencing Mp*murF* lines (sgRNA1)
MurF2_rev	AAGACATTGCAGCTTTGGCG	Genotyping Mp*murF* lines (sgRNA1)

4.3.18. Media and buffers

Media

½ Gamborg medium

Component	**Final Concentration**
Gamborg B5 medium basal salt mixture	1.5 g/l (w/v)
MES	0.5 g/l (w/v)
Sucrose	1 % (w/v)
Agar (Dechefa)	0.8 % (w/v)

M51C medium

Component	**Final concentration**
KNO_3	19.8 µM
NH_4NO_3	5 µM
$MgSO_4$	1.5 µM
$CaCl_2$	2 µM
KH_2PO_4	2 µM
EDTA-Fe(III)	110 nM
Na_2MoO_4	1 µM
$CuSO_4$	100 nM
$CoCl_2$	100 nM
$ZnSO_4$	7 µM
$MnSO_4$	40 µM
H_3BO_3	49 µM
Myo-inositol	555 µM
Nicotinic acid	8 µM
Pyridoxine hydrochloride	5 µM
Thiamine hydrochloride	5 µM
Sucrose	58 mM
L-glutamine	2 mM
Casein hydrolysate	0.1 % w/v

pH adjusted to 5.8

Storage solution

Component	**Final concentration**
Tris-HCl	10 mM
NaCl	100 mM
B-mercaptoethanol	5 mM
EDTA	0.5 mM
Glycerol	50 % w/v
$MgCl_2$	1 mM

pH adjusted to 7.5

4.4. Results

4.4.1. The enzyme MurA is encoded by *Marchantia polymorpha* and is structurally similar to EPSPS

In Chapter 3, I determined that *Marchantia polymorpha* is glyphosate-tolerant but not due to a glyphosate-tolerant EPSPS. To investigate if an EPSPS-like enzyme may contribute to tolerance, I used BLASTP to identify EPSPS-like sequences in the *M. polymorpha* genome. By querying the proteome of *M. polymorpha* with the *A. thaliana* EPSPS amino acid sequence using BLASTP, I identified a similar protein in *M. polymorpha* – UDP-N-acetylglucosamine-3-enopyruvyl transferase (MurA). This enzyme catalyses the first step of peptidoglycan biosynthesis, transferring the enolpyruvyl from phosphoenolpyruvate (PEP) to uridine diphosphate N-acetylglucosamine (UDP-GlcNAc). The structural similarity between MpEPSPS and MpMurA was assessed using AlphaFold predictions and ChimeraX (Fig. 4.2).

The MpEPSPS and MpMurA structural predictions are superficially very similar, likely due to both enzymes catalysing the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) to their respective substrates. Each enzyme consists of 2 domains comprised of α-helices and β-sheets. These domains are joined by a flexible linker region, and in the cleft of the 2 domains forms the active site. The MatchMaker function of USCF ChimeraX was used to quantify the similarity between the two protein structures with the chloroplast transit peptides removed. This function optimally aligns two protein structures by minimizing RMSD to maximise structural similarity. As a comparison, the AlphaFold structure of the EPSPS from *A. thaliana* was also compared to that of the *M. polymorpha* EPSPS. The sequence alignment score for AtEPSPS and MpEPSPS was 1751, with an RMSD of 0.435 Å for 427 pruned atom pairs (0.915 Å for all 439 pairs). In comparison, there was a lower sequence alignment score between MpEPSPS and MpMurA of 582.2, along with a reduced structural similarity indicated by an RMSD of 1.347 Å for 146 pruned atom pairs (14.648 Å for all 451 pairs). However, this still represents a profound similarity in the structures of the 2 proteins, considering their distinct functions.

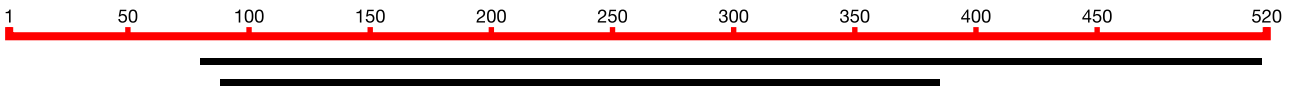
The structural similarity suggests MpMurA could function with EPSPS substrates or bind glyphosate, potentially contributing to glyphosate tolerance in *M. polymorpha*. If MpMurA

contributes to glyphosate tolerance, I hypothesise that this protein is not present in plants that are sensitive to glyphosate. To test this I conducted a BLAST analysis of MurA sequences in genomes from organisms throughout the Archaeplastida.

A

Query= AT1G48860.1_EPSPS_Arabidopsis_thaliana

1 / 1



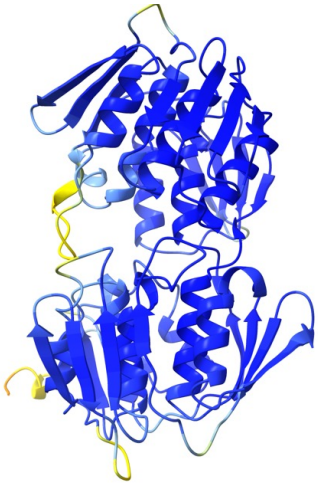
Number	Sequences producing significant alignments		Total score	E value	Length
1.	Mp6g04140.1	(MpEPSPS)	638.65	0.00	527
2.	Mp5g14110.1	(MpMurA)	37.35	1.80×10^{-02}	482

B



MpEPSPS

C



MpMurA

D



Fig. 4.2. The enzyme MpMurA is structurally similar to that of MpEPSPS

A. BLAST result using AtEPSPS as a query against the *Marchantia polymorpha* proteome. Two enzymes are returned, MpEPSPS and MpMurA. MurA catalyses the first step in peptidoglycan biosynthesis.

B. AlphaFold prediction of the EPSPS enzyme from *M. polymorpha*.

C. AlphaFold prediction of the MurA enzyme from *M. polymorpha*.

D. Overlay of the EPSPS and MurA enzymes from *M. polymorpha* produced in USCF ChimeraX using the MatchMaker function.

AlphaFold per residue confidence score (pLDDT) between 0 and 100. Colour denotes confidence in the prediction: Dark blue = Very high (pLDDT > 90); Light blue = Confident (90 > pLDDT > 70); Yellow = Low (70 > pLDDT > 50); Orange = Very low (pLDDT < 50).

4.4.2. EPSPS is conserved throughout the Archaeplastida, but MurA is absent from Angiosperms

Most economically significant glyphosate-controlled weeds are angiosperms. If MurA contributes to glyphosate tolerance in *M. polymorpha*, it should either be absent in angiosperms or exist in a distinct form which does not confer tolerance. To test for the presence of MurA in angiosperms, a BLASTp search was carried out, using the MpMurA as a query, against the proteomes or predicted proteomes of several angiosperm species, as well as species from across the Archaeplastida, *E.coli* and *S. cerevisiae*. Then, by using a supported topology of the Archaeplastida phylogeny, I determined the groups in which MurA was conserved. I also searched the same Archaeplastida species for the presence of EPSPS. A depiction of the Archaeplastida species and clades in which EPSPS and/or MurA are conserved is presented in Fig. 4.2. The MurA amino acid sequences which were retrieved were also aligned and this alignment was used to construct a phylogenetic tree using maximum likelihood statistics (Supplementary Fig. S4.1).

Of all the proteomes or predicted proteomes of Archaeplastida species searched, only that of *Azolla filliculoides* was missing an EPSPS sequence. An EPSPS was encoded in the genomes of all other members of the Archaeplastida searched. An incomplete version of an EPSPS sequence was retrieved from *A. filliculoides*, and this probably represents an artifact of poor quality of proteome coverage of *A. filliculoides*, and not this species lacking a full length EPSPS. On the other hand, MurA was much less conserved throughout the Archaeplastida, with a total of 18 of the 55 species sampled with no MurA enzyme, and 2 having a partial MurA. As hypothesised, no angiosperm species was found to have a full MurA enzyme. However, *Solanum lycopersicum* was found to possess a partial MurA enzyme. This is more dissimilar to MpMurA than most other land plant MurA enzymes, and this is reflected by the long branch of the MurA phylogenetic tree to this node (Supplementary Fig. S4.1). The absence of MurA from all other angiosperm species searched is consistent with my hypothesis that this enzyme could contribute to glyphosate tolerance in *M. polymorpha*. There was a MurA encoding gene in the genomes of most other streptophyte species analysed. Two non-angiosperm streptophyte species in which MurA was absent was an *Isoetes* species and the hornwort *Nothoceros vincentianus*. A MurA was encoded in the

genome of the other lycophyte and hornwort species included in the analysis, suggesting the absence in the aforementioned *Isoetes* species and *Nothoceros vincentianus* were exceptions. There was a partial version of MurA in the moss *Takakia lepidozioide*. Overall, MurA is well conserved among non-angiosperm streptophytes.

MurA is poorly conserved in chlorophyte and rhodophyte algae. Among the sampled chlorophyte algae (11 species), MurA sequences were found in only 3. In the single glaucophyte algae sampled (*Cyanophora paradoxa*), two complete MurA enzymes were identified. Conversely, none of the three sampled rhodophyte algae species had a MurA enzyme, indicating a complete loss of the enzyme in this lineage. In the non-Archaeplastida species analysed, MurA was absent in the yeast *S. cerevisiae* but present in the bacterium *E. coli*.

I hypothesise that MurA contributes to glyphosate tolerance in *M. polymorpha*, and consistent with this, a full MurA enzyme is not encoded in the genomes of angiosperms. A MurA enzyme was found in many other land plant species in this analysis. I hypothesised that these plants would also be tolerant. To test this, I carried out dose response experiments on two other land plant species with a MurA gene.

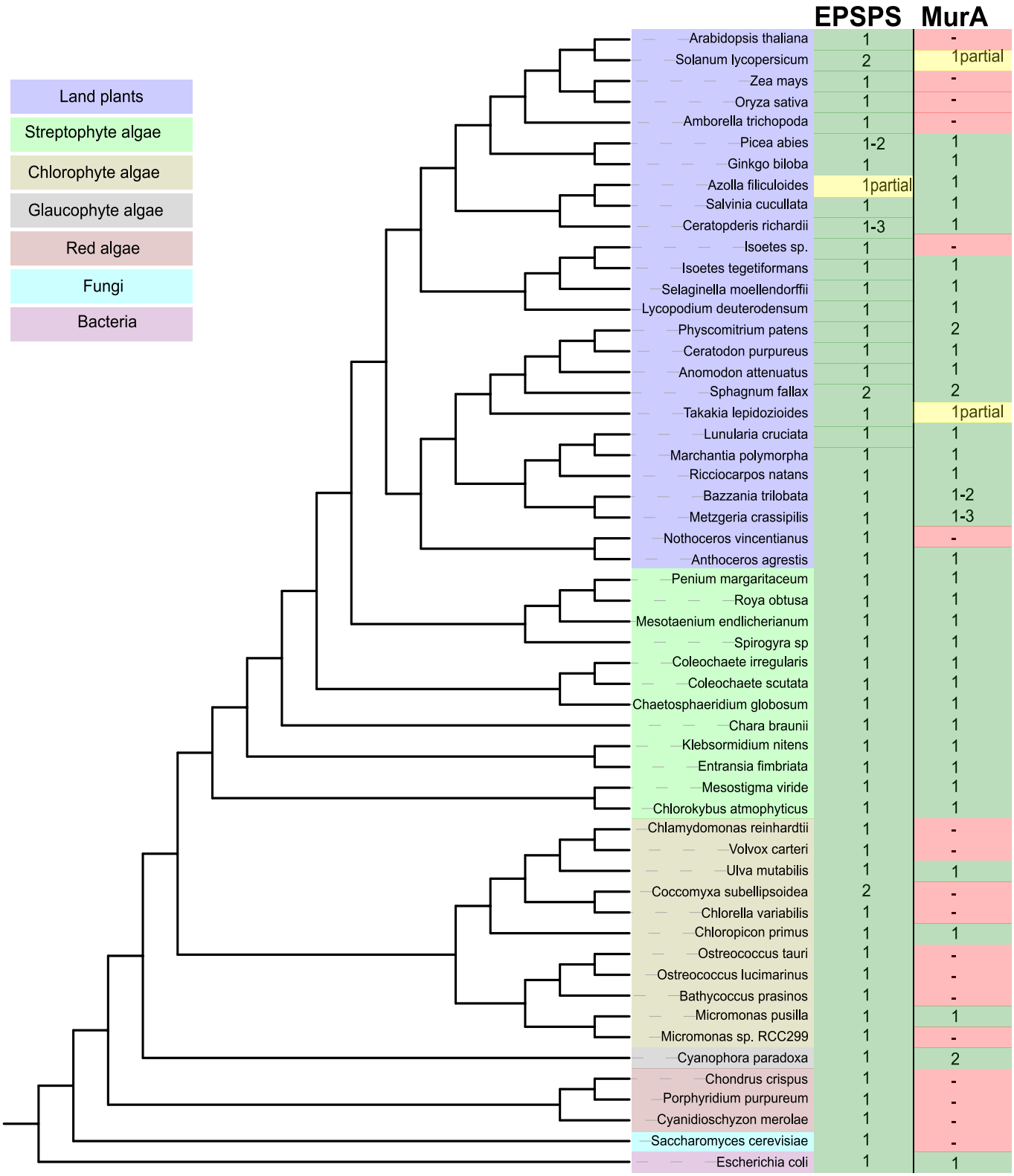


Fig. 4.3. MurA is conserved among many streptophytes, but is absent among angiosperms
Species tree based on: (Carpenter et al., 2019; Harris et al., 2020; Morris et al., 2018). Presence or absence of EPSPS and MurA was searched for in each species using a BLASTp search. The amino acid sequence of each enzyme from *M. polymorpha* was used as a query. Green depicts presence, red absence, and yellow presence of a partial version. The number of versions of enzymes when present is given.

4.4.3. Land plants that encode a MurA enzyme are tolerant to glyphosate

To evaluate the contribution of MurA to glyphosate tolerance, I conducted glyphosate dose-response experiments on additional plant species that encoded a MurA enzyme identified through BLASTP searches. These were the moss *Physcomitrium patens* and the thalloid liverwort *Lunularia cruciata*. As controls, I included *A. thaliana*, which lacks a MurA enzyme and is sensitive to glyphosate, and *M. polymorpha*, which possesses a MurA enzyme and is glyphosate-tolerant (Chapter 3, Section 4.1). All plants were grown for 21-days in diurnal light on media supplemented with glyphosate before imaging (Fig. 4.4). *A. thaliana* was the most sensitive to glyphosate of the plants tested, with 1 µM of glyphosate greatly inhibiting growth. Of the bryophyte species, *M. polymorpha* was the most sensitive to glyphosate. The growth of *M. polymorpha* plants, measured by area after 21 days, was not reduced until a dose of 10 µM. *M. polymorpha* plants were still alive up to and including a dose of 100 µM, however, they showed very little growth at this dose. In contrast, both *L. cruciata* and *P. patens* showed minimal growth reduction across all glyphosate doses. In *L. cruciata*, plants reached the same size at 100 µM of glyphosate as at 0 µM, although a slight visual effect of the 100 µM treatment was observed (Fig. 4.4 B). *P. patens* displayed insensitivity up to 50 µM but exhibited a slight reduction in growth at 100 µM. There was a measurable increase in the area of the two liverwort species at certain glyphosate doses. *M. polymorpha* plants were 1.523-fold larger at 0.5 µM of glyphosate compared to those at 0 µM while *L. cruciata* plants were 1.894-fold larger at 10 µM, a dose lethal to *A. thaliana plants*. This suggests potential compensatory responses in liverwort species leading to growth enhancement at low glyphosate doses.

In summary, both *L. cruciata* and *P. patens*, which encode at least one MurA enzyme, were not only tolerant to glyphosate but more tolerant than *M. polymorpha*. This is consistent with the involvement of MurA in glyphosate tolerance, as all plants encoding a MurA enzyme were more tolerant than *A. thaliana*, which lacks it. However, it is worth noting that the three tested bryophyte species, including *M. polymorpha*, belong to phylogenetically distinct lineages from *A. thaliana*. Therefore, it is formally possible that another shared feature contributes to their common glyphosate tolerance. To test more fully the role of MurA, I investigated how genetic manipulations of the MurA gene influenced glyphosate tolerance in M. polymorpha.

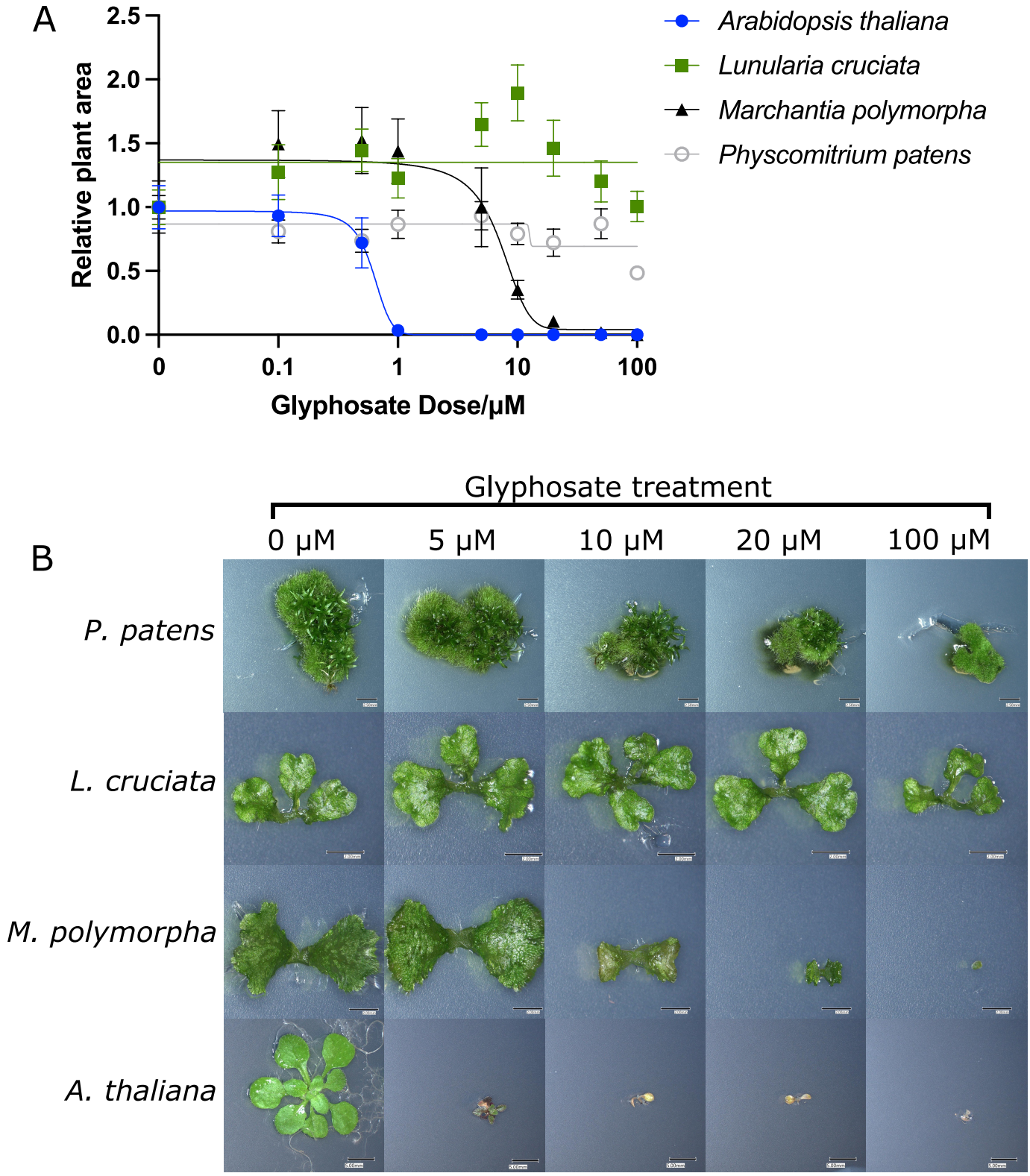


Fig. 4.4. Land plants that possess a MurA enzyme are more tolerant to glyphosate than those that do not

A. Dose response curves of the plant areas of *A. thaliana, M. polymorpha, L. cruciata* and *P. patens* plants grown for 21 days on glyphosate treated plates. The data are fitted with four-parameter log-logistic regression curves. Error bars are ±SE.

B. Brightfield images of individual plants from each species on 5 selected doses of glyphosate.

4.4.4. Overexpression of Mp*MurA* confers increased resistance to glyphosate

I identified MpMurA as a potential contributor to glyphosate tolerance in *M. polymorpha*. To test this hypothesis, I overexpressed MpMurA in *M. polymor*pha by cloning it into an overexpression vector under the control of the cauliflower mosaic virus (CaMV) 35S promoter. The vector was transformed into *M. polymorpha* spores generated from a cross between Tak-1 and Tak-2. As a control, the same vector expressing yellow fluorescent protein (YFP) in place of MpMurA was used. The resulting transformants were evaluated for glyphosate sensitivity (Fig. 4.5). After 14 days growth, plants were imaged on both untreated media and media supplemented with 10 µM of glyphosate – a concentration significantly inhibiting wild-type plant growth. The plant area on glyphosate-treated plates, relative to that on untreated plates was calculated and compared to Tak-1 using a Welch's t-tests (Fig. 4.5 A). Tak-1 was the most glyphosate-tolerant of the wild-type lines, so this was used as the conservative baseline for statistical tests.

The relative area of most plants overexpressing Mp*MurA* grown on a dose of 10 µM was higher than Tak-1 and Tak-2 plants or plants overexpressing YFP (vector controls). This is consistent with overexpression of Mp*MurA* conferring glyphosate resistance. The relative area of five lines was significantly higher than Tak-1 ($p < 0.05$). These lines were 35S:Mp*MurA*_2, 35S:Mp*MurA*_3, 35S:Mp*MurA*_7, 35S:Mp*MurA*_5, and 35S:Mp*MurA*_6. The relative areas of the lines 35S:Mp*MurA*_4, 35S:Mp*MurA*_8, 35S:Mp*MurA*_10, and 35S:Mp*MurA*_11 were also larger than Tak-1 but this difference was not deemed significant ($p > 0.05$).

The 35S:MpMurA_2 line displayed the highest glyphosate resistance among the MpMurA overexpression lines, with a relative plant area of 0.825 (SE = 0.160) at 10 µM glyphosate versus untreated plates. This resistance was statistically significant compared to Tak-1 (relative area = 0.076, SE = 0.03) ($p = 0.0003$), indicating minimal impact of glyphosate treatment at this dose. Furthermore, the plant areas of this line were consistently larger across various glyphosate doses compared to wild-type lines (Fig. 4.5 B). However, both 35S:MpMurA_2 and Tak-1 lines were lethally affected by 50 µM glyphosate, with no surviving plant tissue observed at this concentration in either line.

Overexpression of Mp*MurA* in *M. polymorpha* significantly increases resistance to glyphosate in five of the nine lines tested. This increase in resistance is consistent with MurA being involved in glyphosate tolerance in *M. polymorpha*.

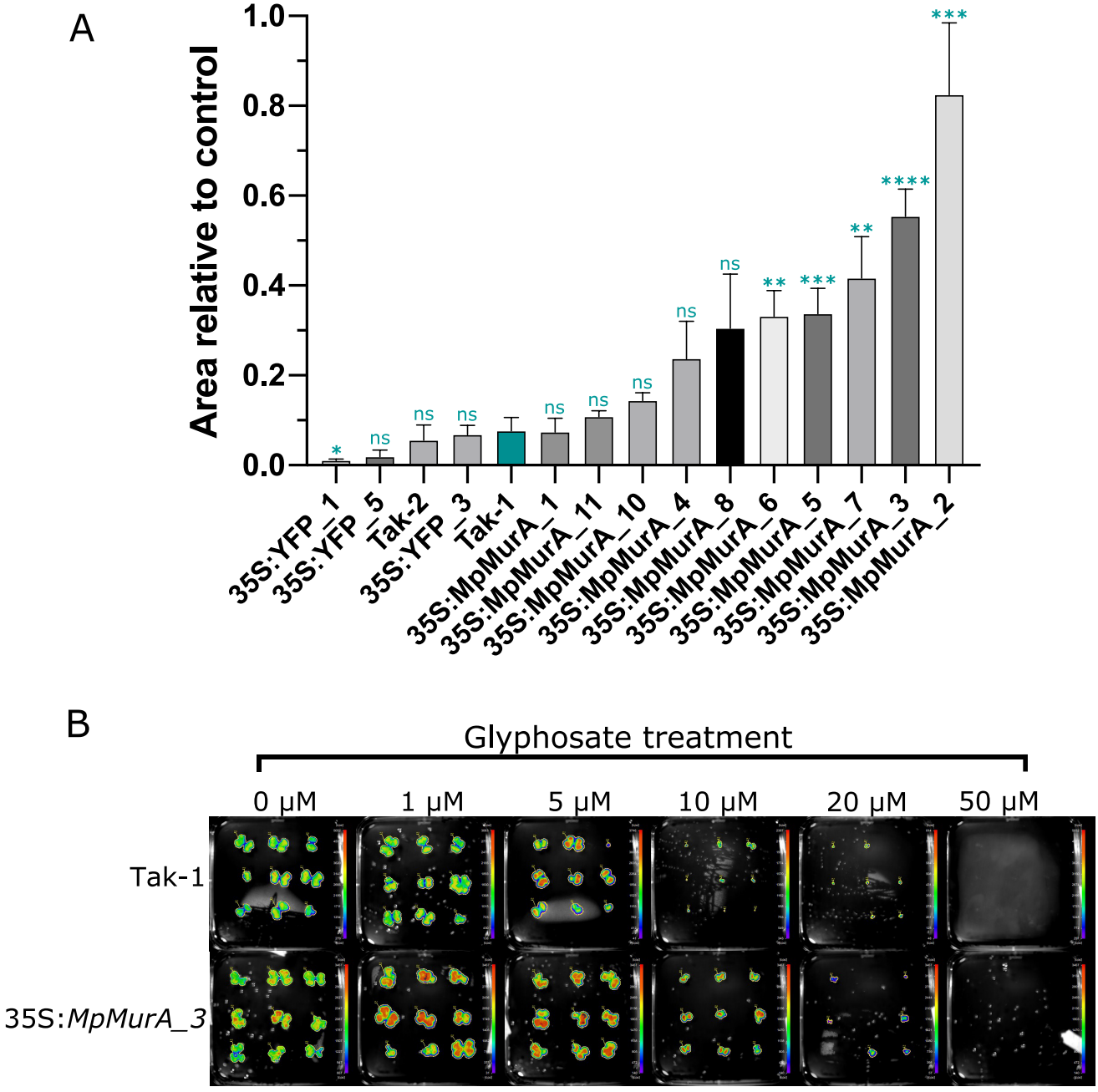


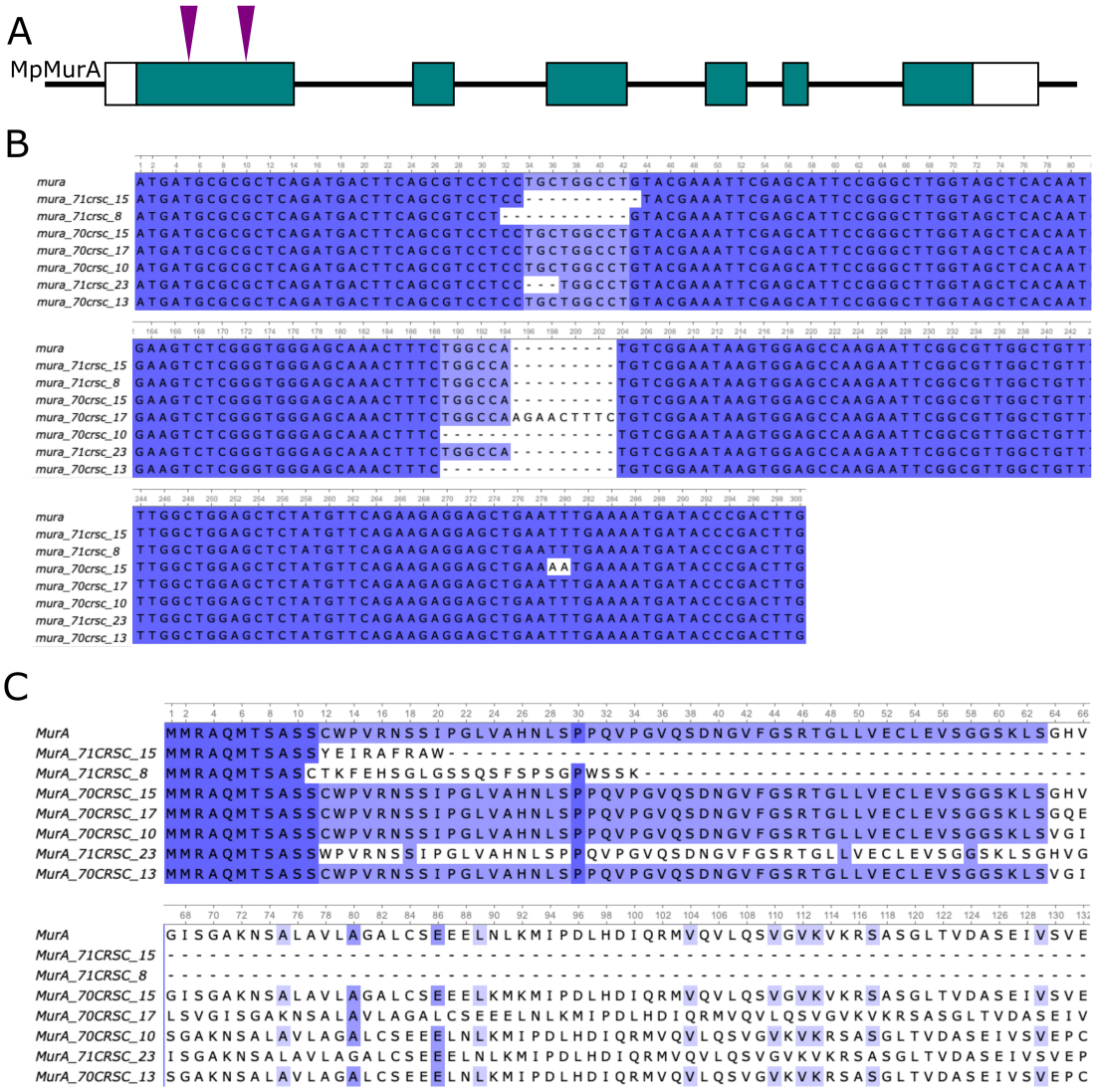
Fig. 4.5. Overexpression of Mp*MurA* in *M. polymorpha* confers an increase in resistance to glyphosate in five out of nine lines

A. *M. polymorpha* wild-type and Mp*MurA* overexpression lines were grown on 0 µM and 10 µM of glyphosate. The auto-fluorescing areas were measured and the ratio between treated and untreated plants calculated and plotted. Error bars are ±SE. Welch's t-tests were performed using Tak-1 area relative to control as the baseline *=$p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, n.s. = not significant.

B. Images of the chlorophyll auto-fluorescing area of Tak-1 and 35S:Mp*MurA*_3 on 6 selected doses of glyphosate.

4.4.5. Seven predicted Mp*murA* loss-of-function mutants were generated using CRISPR-Cas9 mutagenesis

To independently test the effect of MpMurA function on glyphosate tolerance in *M. polymorpha,* I generated Mp*murA* loss-of-function mutants using the CRISPR-Cas9 mutagenesis system. I designed two guide RNAs to target the first exon of the Mp*MurA* gene, with the aim of inducing large mutations or frameshifts which would lead to strong loss-of-function of the expressed protein (Fig. 4.6 A). Guide RNAs were cloned into vectors which contained the coding sequence for the Cas9 enzyme. These vectors were transformed into wild-type *M. polymorpha* spores from a cross between Tak-1 and Tak-2 using *A. tumefaciens* (Section 4.3.8). Positive transformants were genotyped and a total of seven Mp*murA* mutants were found, with a variety of mutations in the nucleotide sequence around the two guide RNA binding sites (Fig. 4.6 B). To determine the amino acid sequence of the expressed MurA enzyme in Mp*murA* mutants, a predicted translation of the coding sequence of the Mp*murA* genes was produced (Fig. 4.6 C). I identified two mutants in which the mutation induced by CRISPR-Cas9 would result in a frame shift, causing the predicted expression of a truncated version of the protein. These mutants (*murA_71CRSC_15* and *murA_70CRSC_8*) are predicted to express a very strong loss-of-function MurA enzyme. The remaining 5 mutants had indels or substitutions of few amino acids in the expressed protein, and so these were predicted to be putative weak loss-of-function mutants. The nature of the mutations predicted to be induced by CRISPR-Cas9 mutagenesis are summarised in Fig. 4.6 D.



D

Mp*murA*^{lof} line	Nucleotide sequence change	Amino acid sequence change
71CRSC_15	10bp deletion	Truncation 9 aa downstream
71CRSC_8	11bp deletion	Truncation 24 aa downstream
70CRSC_15	2bp substitution	Asn>Lys, Leu>Met substitutions
70CRSC_17	9bp insertion	His>Gln substitution, insertion of Glu, Leu, Ser
70CRSC_10	6bp deletion	Deletion of Gly, His
71CRSC_23	3bp deletion	Deletion of Cys
70CRSC_13	6bp deletion	Deletion of Gly, His

Fig. 4.6. Generation of predicted loss-of-function mutation in Mp*MurA* by CRISPR/Cas9 mutagenesis

A. Mp*murA* mutant lines were generated using the CRISPR-Cas9 mutagenesis system with two guide RNAs designed to target the first exon of the Mp*MurA* gene carried by a vector containing the Cas9 gene which were transformed into *M. polymorpha* spores.

B. Positive transformants were genotyped using Sanger sequencing to determine the nature of the induced mutation on the coding sequence of Mp*MurA*. Sequences were aligned to the wild-type reference (MpTak1v6.1) using MAFFT via the L-INS-i algorithm.

C. The predicted expressed amino acid sequences of MurA proteins in Mp*murA* mutants were aligned to the reference using MAFFT via the L-INS-i algorithm.

D. A summary of the nature of the mutations induced in Mp*MurA* induced by CRISPR-Cas9 mutagenesis.

Mutations in peptidoglycan biosynthesis genes have been reported to result in a distinct macrochloroplast phenotype in plants (Homi et al., 2009; Machida et al., 2006). To test if my Mp*murA* predicted loss-of-function mutants had the same phenotype and could therefore be confirmed as MurA loss-of-function mutants, I produced brightfield images of three-day old gemmae from Tak-1, *murA_70crsc_10*, *murA_71crsc_8*, and *murA_71crsc_15* in which the chloroplasts were visible (Fig. 4.7). *murA_70crsc_10* (Fig. 4.7 D-F) is predicted to be a weak loss-of-function mutant, with a predicted 6bp deletion in the coding sequence of Mp*MurA* leading to a deletion of a glycine and histidine from the amino acid sequence of the expressed protein. The other Mp*murA* mutants used in this analysis are both predicted to be strong loss-of-function mutants. *murA_71crsc_8* (Fig. 4.7 G-I) is predicted to express a version of the MurA protein which is truncated 24 amino acids downstream of the mutation site while *murA_71crsc_15* (Fig. 4.7 J-L) is predicted to express an even shorter version of the MurA enzyme, with a truncation 9 amino acids downstream. Tak-1 (Fig. 4.7 A-C) was used as a control for this analysis. In all three Mp*murA* mutant lines, the chloroplasts were visibly much larger than in the Tak-1 gemmalings (Fig. 4.7). The area of all three of these mutants was smaller than Tak-1 after 3 days of growth. However, only in *murA_71crsc_15* were the gemmalings misshapen in any way, losing the lobed form of the Tak-1, and other mutant gemmalings.

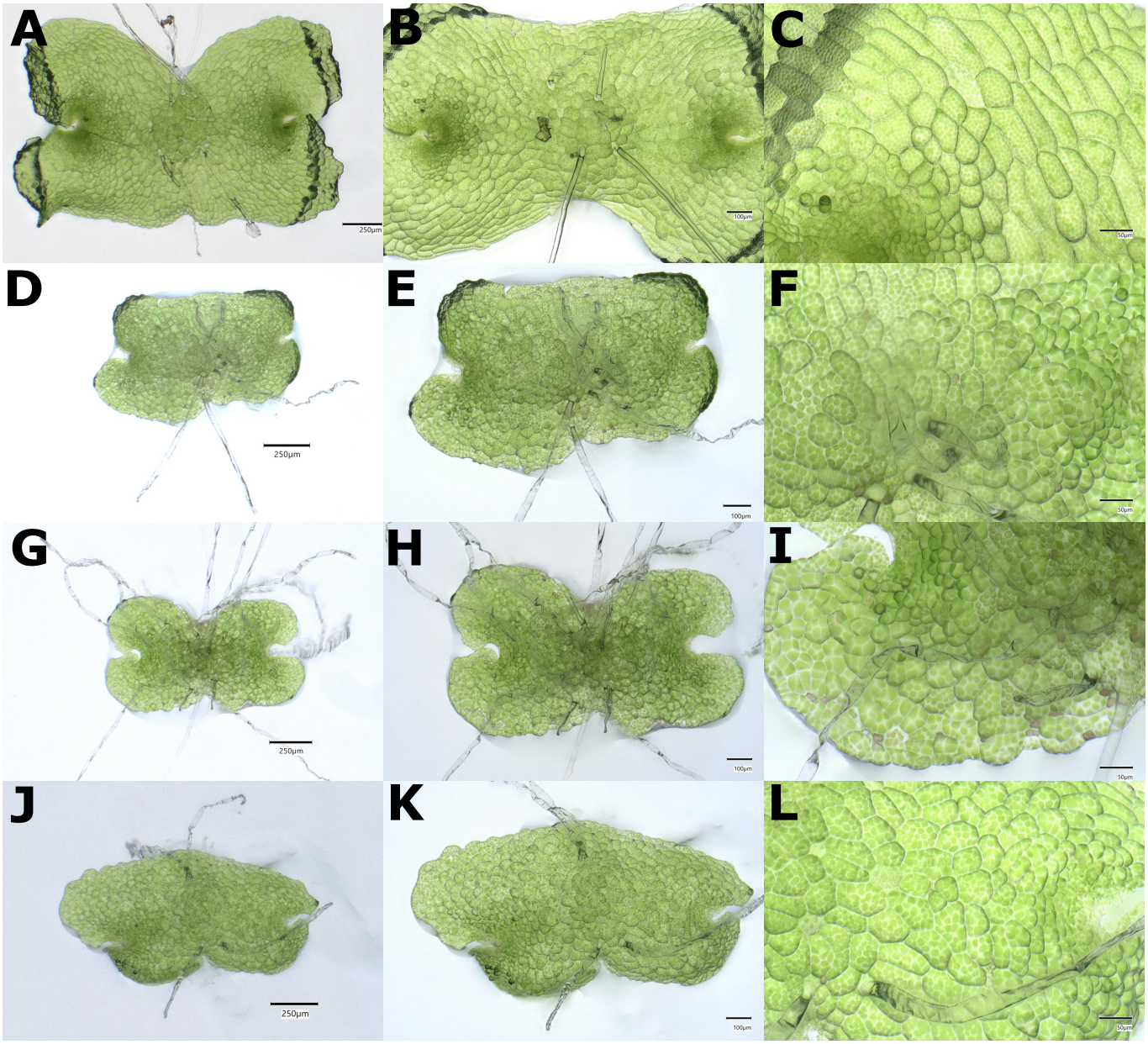


Fig. 4.7. Mp*murA* loss-of-function predicted mutants have a distinct macrochloroplast phenotype

The gemmaling phenotypes of three Mp*murA* loss-of-function lines and wild-type *M. polymorpha* were assessed through imaging with a Keyence VHX-7000 microscope: (A-C) Tak-1; (D-F) *murA_70crsc_10*; (G-I) *murA_71crsc_8*; (J-K) *murA_71crsc_15*. A macrochloroplast phenotype was identified in all three mutant lines and morphological differences at the whole gemmaling level were also observed.

The images of chloroplasts in the cells of gemmalings taken using transmitted light were then used to quantify chloroplast size differences in Mp*murA* mutants (Fig. 4.8). The image analysis software FIJI was used to measure various chloroplast and cell dimensions. Over 40 chloroplasts and 20 cells were measured for each mutant and wild-type line, sampled from the flattest and most mature tissue in multiple gemmalings. Chloroplast measurements were statistically compared to those in Tak-1 gemmalings using a one-way ANOVA and Dunnett's multiple comparison test.

The mean longest and shortest axes of the chloroplasts in all three predicted Mp*murA* loss-of-function mutants were significantly larger than those in the wild-type (adjusted p-values < 0.0001). Among them, *murA_71crsc_8* exhibited the largest long and short chloroplast axes (Fig. 4.8 A and B). Additionally, the long-to-short axis ratio of chloroplasts was greater in all three Mp*murA* mutants compared to the wild-type ($p < 0.025$), indicating chloroplast elongation in the mutants, as observed in Fig. 4.7. There were fewer chloroplasts in Mp*murA* gemmaling cells compared to Tak-1 gemmaling cells ($p < 0.0001$) (Fig. 4.8 D), and the cell size in Mp*murA* loss-of-function mutants was smaller than in wild-type ($p < 0.001$) (Fig. 4.8 E). The reduced number of chloroplasts per cell was not attributable to smaller cell size, as the number of chloroplasts per unit of cell area was significantly lower in the Mp*murA* predicted loss-of-function mutants ($p < 0.0001$) (Fig. 4.8 F).

In summary, these data are consistent with the Mp*murA* predicted loss-of-function mutants generated (outlined in Fig. 4.6) expressing defective versions of the MpMurA enzyme. Since both large and small amino acid changes in the expressed protein lead to significant alterations in chloroplast morphology and abundance, all the MpmurA predicted loss-of-function mutants can be considered confirmed loss-of-function mutants.

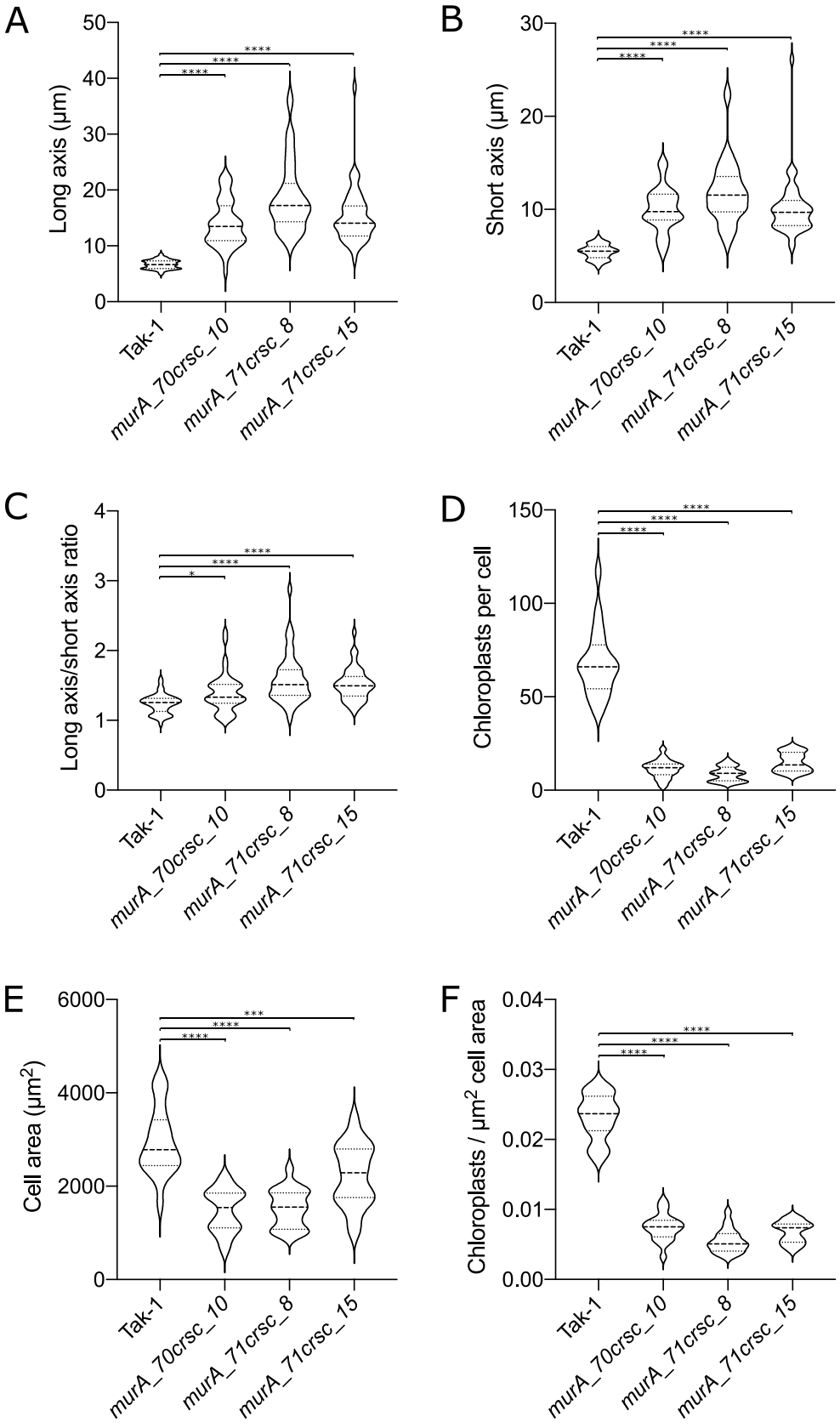


Fig. 4.8. Mp*murA* predicted loss-of-function mutants have fewer, larger chloroplasts than wild-type *M. polymorpha* plants

Violin plots of various chloroplast and cell measurements of wild-type (Tak-1) and Mp*murA* predicted loss-of-function mutants were calculated using the image analysis software FIJI. The chloroplast longest axis (A), shortest axis (B), long/short axis ratio (C), chloroplast abundance per cell (D), cell area (E) and chloroplast abundance per unit area of cell (F) were measured or calculated. A one-way ANOVA and Dunnett's multiple comparison test was performed, to compare values in Mp*murA* mutants to Tak-1, *=p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001, n.s. = not significant.

4.4.6. Predicted Mp*murA* loss-of-function mutants are more sensitive to glyphosate than wild-type plants

I generated seven Mp*murA* loss-of-function mutants using CRISPR-Cas9 mutagenesis, which were predicted to lack the expression of functional MurA enzyme. The observed macrochloroplast phenotype in these mutants (Section 4.4.5) is consistent with MurA loss-of-function. To test if MurA contributes to glyphosate tolerance in *M. polymorpha*, I assessed the sensitivity of Mp*murA* loss-of-function mutants to glyphosate compared to wild-type plants (Fig. 4.9). Both wild-type and mutant plants were grown for 14 days on media with 5 µM glyphosate (a dose that doesn't significantly inhibit wild-type growth) and untreated media before imaging. Measured plant area was then used to calculate the ratio between the area on treated and untreated media (Fig. 4.9 A).

Tak-2 is more sensitive to glyphosate than Tak-1 and so this was used as a conservative baseline for statistical comparison between Mp*murA* lines and wild-type lines. All Mp*murA* plants were more sensitive to glyphosate than both Tak-1 and Tak-2 (Fig. 4.9 A). These differences were all statistically significant (Welch's t-test, $p < 0.05$). All of the Mp*murA* lines tested were at least twice as sensitive to 5 µM of glyphosate as the most sensitive wild-type accession, Tak-2. The line *murA_71CRCS_23* had the largest area relative to control of all the mutant plants with 0.167 (SE = ±0.032). The area relative to control of Tak-2 was 0.355 (SE = ±0.046).

The most sensitive Mp*murA* line was *murA_70CRSC_15*, with an area relative to control of just 0.019 (SE = 0.014), indicating near-lethality at 5 µM glyphosate. Almost no living plant tissue was detected at this glyphosate concentration in this line, with a substantial amount being detected in both the wild-type lines (Fig. 4.9 B). Glyphosate concentrations higher than 5 µM were lethal to *murA_70CRSC_15* plants. The varying sensitivity of the Mp*murA* lines may stem from different mutation types introduced or potential genetic variability resulting from the cross between Tak-1 and Tak-2 plants.

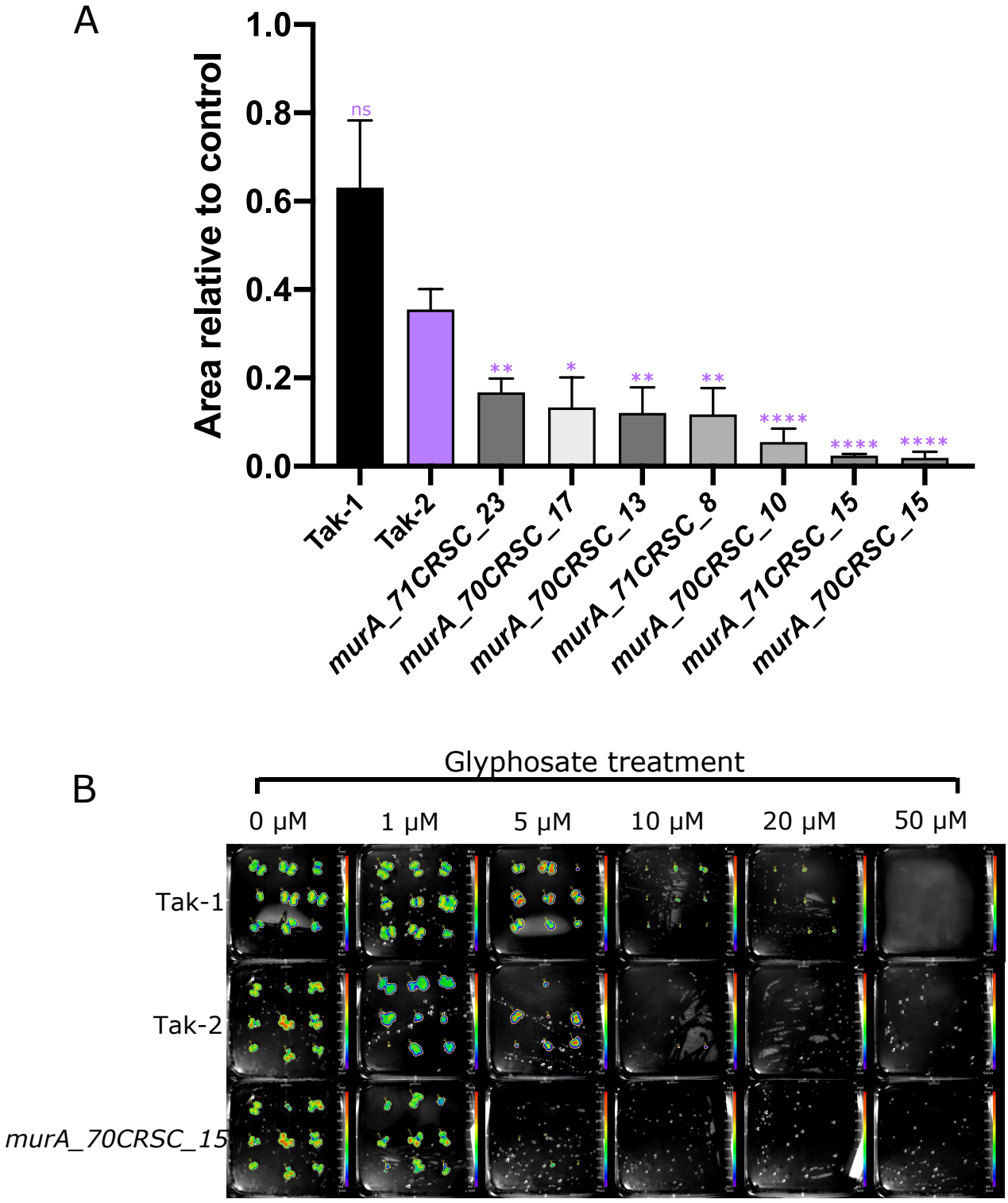
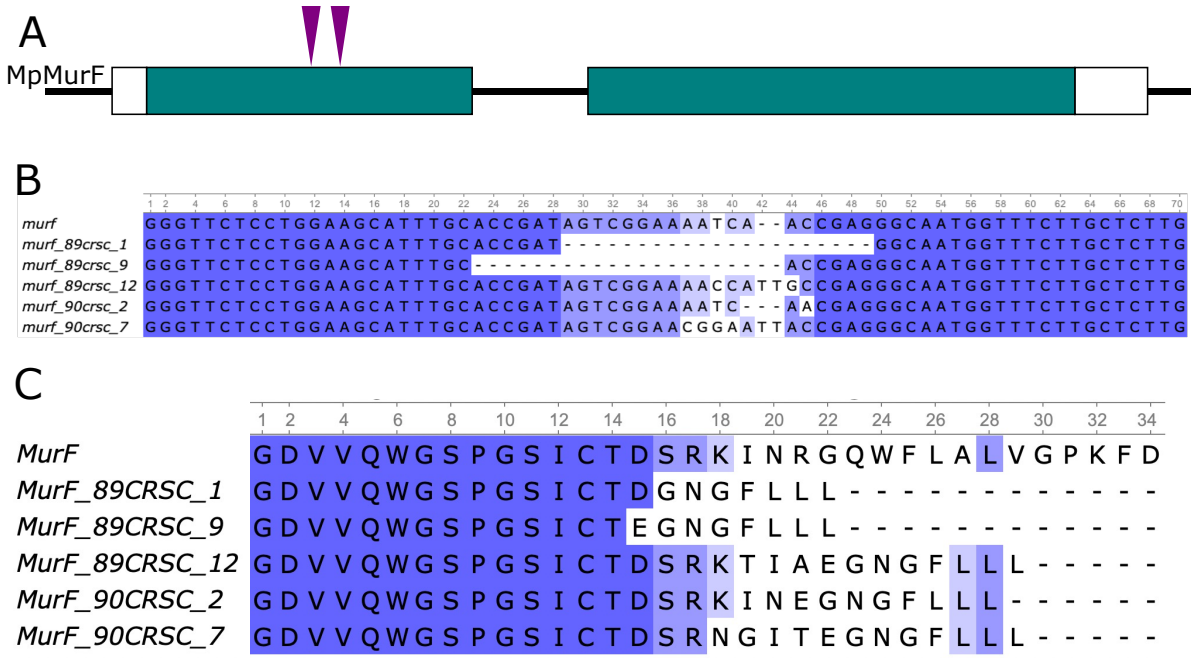


Fig. 4.9. Predicted Mp*murA* loss-of-function mutants are more sensitive to glyphosate
A. *M. polymorpha* wild-type and Mp*MurA* loss-of-function lines were grown on 0 µM and 5 µM of glyphosate. The auto-fluorescing areas were measured and the ratio between treated and untreated plants calculated and plotted. Error bars are ±SE. Welch's t-tests were performed using Tak-2 area relative to control as the baseline *=$p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, n.s. = not significant.
B. Images of the chlorophyll auto-fluorescing area of Tak-1, Tak-2, and the *murA_70CRSC_15* line on 6 selected doses of glyphosate.

4.4.7. Five predicted Mp*murF* loss-of-function mutants were generated using CRISPR-Cas9 mutagenesis

MurA catalyses the initial step of peptidoglycan biosynthesis. It is formally possible that glyphosate tolerance in *M. polymorpha* is due to alterations in peptidoglycan biosynthesis rather than MurA itself. Under this hypothesis, perturbing other stages of peptidoglycan biosynthesis, either genetically or chemically, would affect glyphosate sensitivity in *M. polymorpha*. To test this, I generated Mp*murF* loss-of-function plants using CRISPR-Cas9. MurF is another essential enzyme in peptidoglycan biosynthesis, specifically in forming UDP-MurNAc-pentapeptide by adding D-alanyl-D-alanine to an intermediate molecule. Two guide RNAs targeting the first exon of the MpMurF gene were designed to induce significant loss-of-function mutations (Fig. 4.10 A). The CRISPR-Cas9 procedure for MpMurF followed the same steps as described in Section 4.4.5 for MpMurA. I identified five putative Mp*murF* loss-of-function mutants with various mutations near the guide RNA binding site (Fig. 4.10 B). These mutations resulted in frameshifts and premature stop codons in the coding sequence, predicting the expression of truncated proteins (Fig. 4.10 C). Consequently, all five generated Mp*murF* mutants are expected to be strong loss-of-function mutants. A summary of the predicted mutations in the Mp*murF* lines can be found in Fig. 4.10 D.



D

Mp*murF*^{lof} line	Nucleotide sequence change	Amino acid sequence change
89CRSC_1	19bp deletion	Truncation 7 aa downstream
89CRSC_9	19bp deletion	Truncation 8 aa downstream
89CRSC_12	2bp insertion, 2bp substitution	Truncation 11 aa downstream
90CRSC_2	1bp deletion	Truncation 10 aa downstream
90CRSC_7	2bp insertion, 4bp substitution	Truncation 12 aa downstream

Fig. 4.10. Generation of predicted loss-of-function mutation in Mp*MurF* by CRISPR/Cas9 mutagenesis

A. Mp*murF* mutant lines were generated using the CRISPR-Cas9 mutagenesis system with two guide RNAs designed to target the first exon of the Mp*MurF* gene carried by a vector containing the Cas9 gene which were transformed into *M. polymorpha* spores.

B. Positive transformants were genotyped using Sanger sequencing to determine the nature of the induced mutation on the coding sequence of Mp*MurF*. Sequences were aligned the wild-type reference (MpTak1v6.1) using MAFFT via the L-INS-i algorithm.

C. The predicted expressed amino acid sequences of MurF proteins in Mp*murF* mutants were aligned to the reference using MAFFT via the L-INS-i algorithm.

D. A summary of the mutations induced in Mp*MurF* induced by CRISPR-Cas9 mutagenesis.

To test whether mutations in the Mp*MurF* gene, resulting in predicted loss of protein function, induced the macrochloroplast phenotype associated with peptidoglycan biosynthesis disruption, I produced brightfield images of three-day-old gemmae from Tak-1, *murF_89crsc_12*, and *murF_89crsc_9* in which chloroplasts were visible (Fig. 4.11). Both *murF_89crsc_12* and *murF_89crsc_9* (Fig. 4.11 D-F and G-I, respectively) are strong loss-of-function mutants with mutations in the coding sequences of Mp*MurF*, leading to severe protein truncation and a predicted loss of function (Fig. 4.10 D). Tak-1 served as a control in this analysis (Fig. 4.11 A-C). Chloroplasts in Mp*murF* loss-of-function gemmae were visibly larger than those in Tak-1 plants. The thallus area of *murF_89crsc_12* was smaller after three days of growth compared to Tak-1, and these mutants exhibited more thallus lobes due to increased bifurcations. Additionally, the central sections of *murF_89crsc_12* gemmae appeared paler and translucent. There was less of a reduction in the area of murF_89crsc_9 gemmalings compared to Tak-1. These findings suggest that disrupting MurF leads to a macrochloroplast phenotype, supporting the hypothesis that MurF is required for chloroplast division and peptidoglycan biosynthesis.

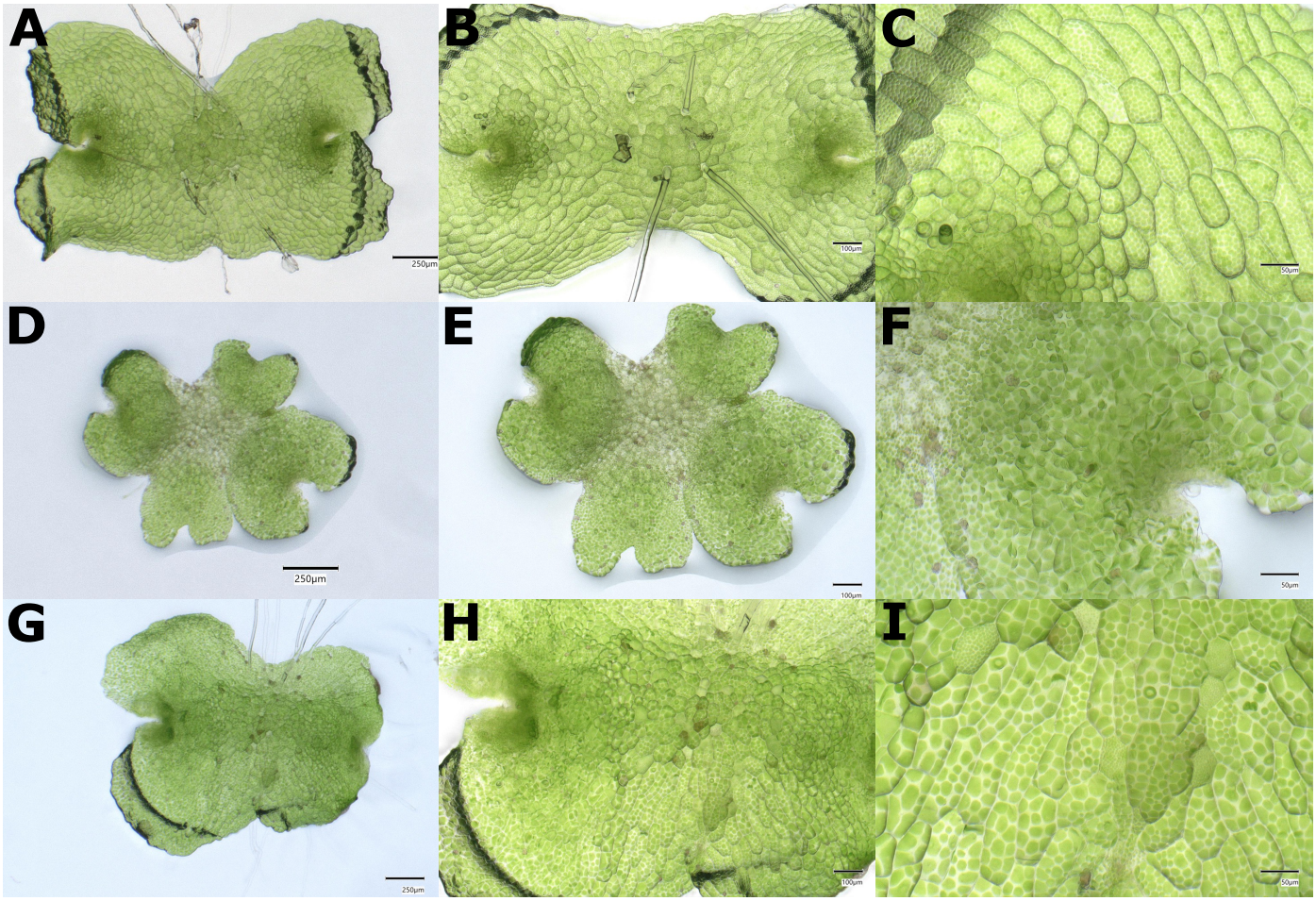


Fig. 4.11. Mp*murF* predicted loss-of-function mutants have a distinct macrochloroplast phenotype

The gemmaling phenotypes of two Mp*murF* loss-of-function lines and wild-type *M. polymorpha* were assessed through imaging with a Keyence VHX-7000 microscope: (A-C) Tak-1, (D-F) *murF_89crsc_12*, and (G-I) *murF_89crsc_9*. A macrochloroplast phenotype was identified in the mutant lines and morphological differences at the whole gemmaling level were also observed.

Images of Tak-1, *murF_89crsc_12*, and *murF_89crsc_9* gemmae taken using transmitted light were used to quantify chloroplast size differences between Mp*murF* loss-of-function mutants and Tak-1 (Fig. 4.12). The analysis methodology was the same as that conducted for Mp*murA* loss-of-function chloroplasts in Section 4.4.5. Chloroplast measurements in gemmae from Mp*murF* loss-of-function lines were compared to those in Tak-1 gemmae using a one-way ANOVA and Dunnett's multiple comparison test. Both the long and short axes of chloroplasts in Mp*murF* loss-of-function gemmae exceeded those in Tak-1 gemmae (adjusted $p < 0.0001$) (Fig. 4.12 A and B). Furthermore, the long-to-short axis ratio of chloroplasts in Mp*murF* loss-of-function gemmae was higher than in Tak-1 gemmae ($p < 0.0001$) (Fig. 4.12 C), indicating less rounded chloroplasts in the mutants. Chloroplast abundance in *murF_89crsc_12* gemmae was reduced more than 8-fold compared to Tak-1 ($p < 0.0001$), and *murF_89crsc_9* exhibited nearly a 7-fold reduction ($p < 0.0001$) (Fig. 4.12 D). The mean cell area in *murF_89crsc_12* gemmae was smaller than in Tak-1 ($p < 0.0001$), while no reduction in mean cell area was observed in *murF_89crsc_9* mutants ($p = 0.8615$). The decreased chloroplast abundance in *murF_89crsc_12* gemmae was not due to reduced cell area compared to Tak-1 gemmae ($p < 0.0001$) (Fig. 4.12 E), as the chloroplasts per unit cell area in *murF_89crsc_12* gemmae remained approximately 2.6-fold lower than in Tak-1 gemmae (Fig. 4.12 F). These findings are consistent with Mp*murF* loss-of-function gemmae expressing a complete loss-of-function MpMurF enzyme, leading to peptidoglycan biosynthesis disruption and chloroplast division defects.

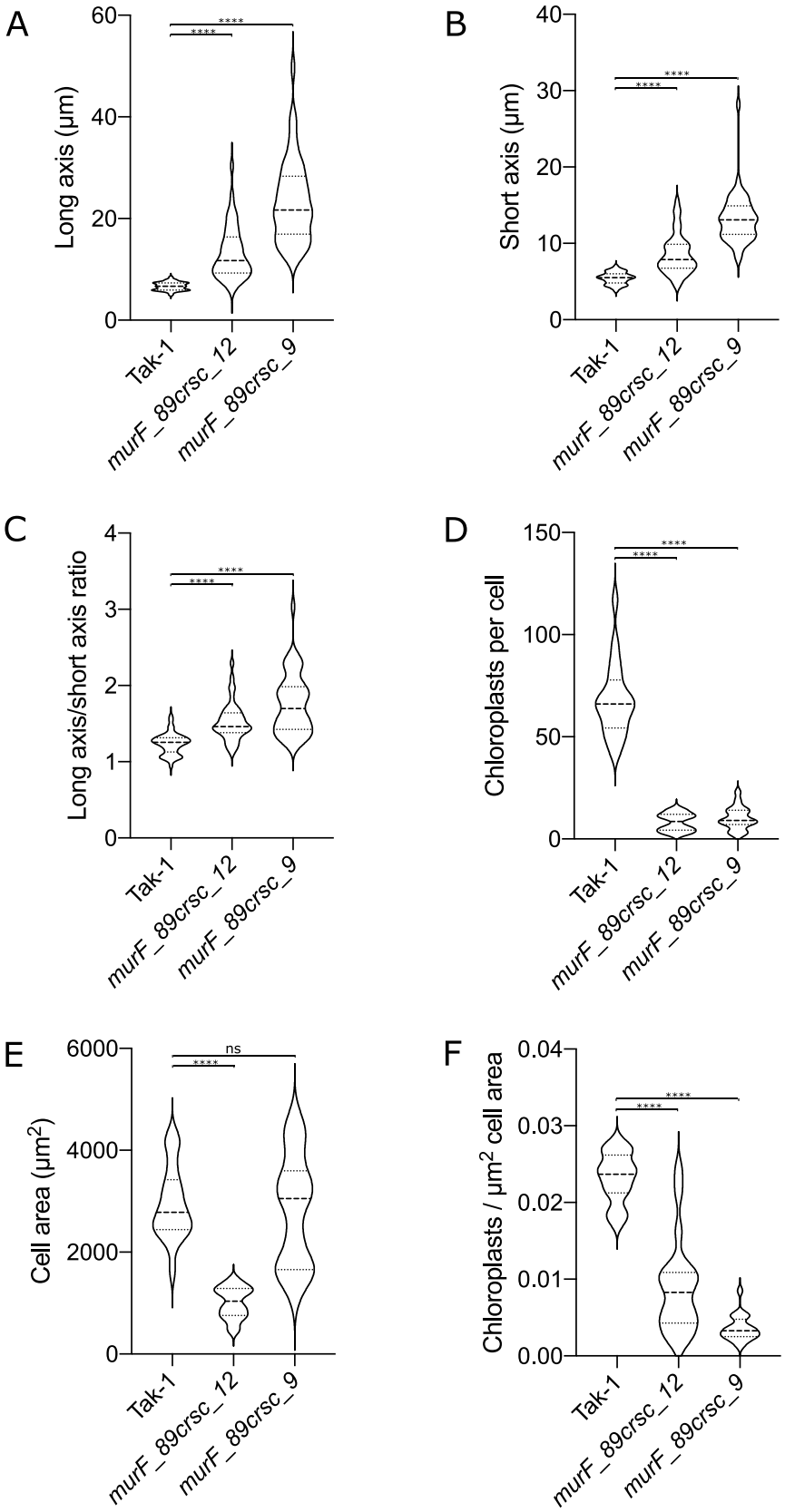


Fig. 4.12. Mp*murF* predicted loss-of-function mutants have fewer, larger chloroplasts than wild-type *M. polymorpha* plants

Violin plots of various chloroplast and cell measurements of wild-type (Tak-1) and two Mp*murF* predicted loss-of-function mutants were calculated using the image analysis software FIJI. The chloroplast longest axis (A), shortest axis (B), long/short axis ratio (C), chloroplast abundance per cell (D), cell area (E) and chloroplast abundance per unit area of cell (F) were measured or calculated. A one-way ANOVA and Dunnett's multiple comparison test was performed, to compare values in Mp*murA* mutants to Tak-1, 0.01, **** = p < 0.0001, n.s. = not significant.

4.4.8. Predicted Mp*murF* loss-of-function mutants are not more sensitive to glyphosate than wild-type plants

To determine if MpMurA or more generally peptidoglycan biosynthesis was responsible for glyphosate tolerance in *M. polymorpha*, I tested if the Mp*murF* loss-of-function mutants I generated in Section 4.4.7 were more sensitive to glyphosate than wild-type plants (Fig. 4.13). MpMurF catalyses a different step of peptidoglycan biosynthesis than MpMurA and so if Mp*murF* predicted loss-of-function mutants were more sensitive to glyphosate, it would be consistent with peptidoglycan biosynthesis, not MpMurA itself, contributing to glyphosate tolerance in *M. polymorpha*.

To test this hypothesis, both wild-type and Mp*murF* mutant plants were grown on media supplemented with 5 µM glyphosate and untreated media for 14 days. The living areas of the plants were measured, and the ratio calculated between the area on treated and untreated media (Fig. 4.13 A). Statistical comparisons were made between these ratios and those calculated for both Tak-1 and Tak-2 plants. Overall, Mp*murF* predicted loss-of-function mutants were not more sensitive or resistant to glyphosate than Tak-1 and Tak-2 plants (Welch's t-test, $p > 0.05$). However, one Mp*murF* predicted loss-of-function line, *murF_89crsc_12*, was significantly larger compared to control than the Tak-2 accession (Welch's t-test, $p = 0.0018$). This suggests this line is slightly more resistant to a dose of 5 µM glyphosate than Tak-2.

Plant area was also measured for eight other doses of glyphosate for *murF_89crsc_12* and *murF_90crsc_2*, and dose response curves were generated by fitting a four-parameter log-logistic equation to these data (Fig. 4.13 B). This was to test if there were differences in sensitivity to glyphosate of Mp*murF* predicted loss-of-function lines at a wider range of glyphosate doses. However, both lines tested exhibited a similar dose-response relationship to glyphosate as Tak-1 and Tak-2 across the range of doses tested. This is consistent with Mp*murF* predicted loss-of-function mutants and wild-type (Tak-1 and Tak-2) being equally sensitive to glyphosate.

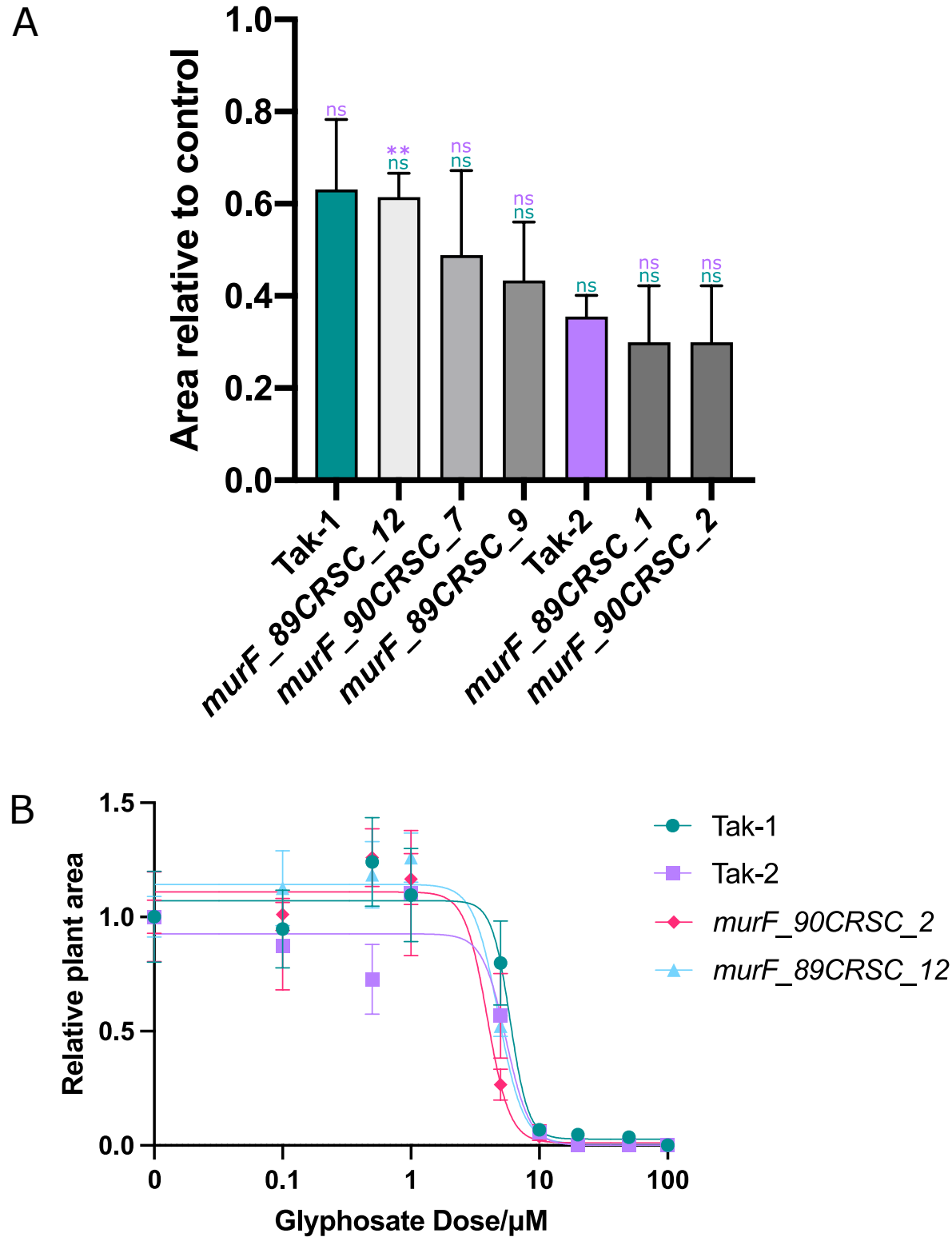


Fig. 4.13. Mp*murF* predicted loss-of-function mutants are not more sensitive to glyphosate
A. *M. polymorpha* wild-type and Mp*murF* loss-of-function lines were grown on 0 µM and 5 µM of glyphosate. The auto-fluorescing areas were measured and the ratio between treated and untreated plants calculated and plotted. Error bars are ±SE. Welch's t-tests were performed using Tak-1 and Tak-2 area relative to control as baselines *=p < 0.05, ** = p < 0.01, n.s. = not significant.
B. Dose response curves of the plant areas of *M. polymorpha* wild-type and Mp*murF* loss-of-function plants grown for 14 days on glyphosate treated plates. The data are fitted with four-parameter log-logistic regression curves. Error bars are ±SE.

4.4.9. Treatment of *M. polymorpha* with some peptidoglycan biosynthesis inhibiting antibiotics leads to a macrochloroplast phenotype

To independently distinguish between peptidoglycan biosynthesis, or the MurA protein being the cause of glyphosate tolerance in *M. polymorpha*, I manipulated peptidoglycan biosynthesis using the antibiotics cefotaxime, ampicillin, fosfomycin and vancomycin. These antibiotics inhibit different stages of peptidoglycan biosynthesis. From the literature, and my experiments in Section 4.4.5 and Section 4.4.7, it is known that perturbation of peptidoglycan biosynthesis in plants can lead to a macrochloroplast phenotype (Dowson et al., 2022; Matsumoto et al., 2012). Therefore, to verify that treatment with the five antibiotics inhibited peptidoglycan biosynthesis, I imaged antibiotic-treated *M. polymorpha* gemmalings 3 days after treatment with antibiotics and assessed and quantified their chloroplast phenotypes, comparing these to untreated control plants (Fig. 4.14 and Fig. 4.15).

Plants treated with cefotaxime or ampicillin developed machrochloroplasty, while treatment of plants with fosfomycin or vancomycin did not alter chloroplast size. These data indicate that disruption of peptidoglycan synthesis with cefotaxime or ampicillin modify the peptidoglycan wall of chloroplasts leading to defects in chloroplast division and macrochloroplasty. Cefotaxime, which inhibits the transpeptidation of peptidoglycan by binding to penicillin-binding proteins (PBPs), causes gemmalings to have visibly larger chloroplasts than untreated Tak-1 gemmalings (Fig. 4.14 F). However, cefotaxime does not impact Tak-1 gemmalings at the whole-plant and cell level (Fig. 4.14 D and E). Treatment with ampicillin also led to visibly enlarged chloroplasts but caused no change in gemmaling or cell size or shape (Fig. 4.14 J-L). Like cefotaxime, ampicillin inhibits peptidoglycan biosynthesis by binding to PBPs. Fosfomycin treatment did not cause a visible change in chloroplast size or cell size and shape (Fig. 4.14 I). Treatment with fosfomycin also resulted in no major change to gemmaling size or shape (Fig. 4.14 G and H). Fosfomycin inhibits peptidoglycan biosynthesis by inhibiting MurA. Vancomycin also did not cause any visible change to chloroplast size but did affect Tak-1 gemmaling size, causing a reduction in the growth of Tak-1 gemmalings compared to untreated Tak-1 gemmalings (Fig. 4.14 M-N). Vancomycin works by preventing certain peptides from being incorporated into the

peptidoglycan matrix. Vancomycin also affects RNA synthesis (Lundstrom and Sobel, 1995), which may explain the influence of this antibiotic on gemmaling size but not chloroplast size.

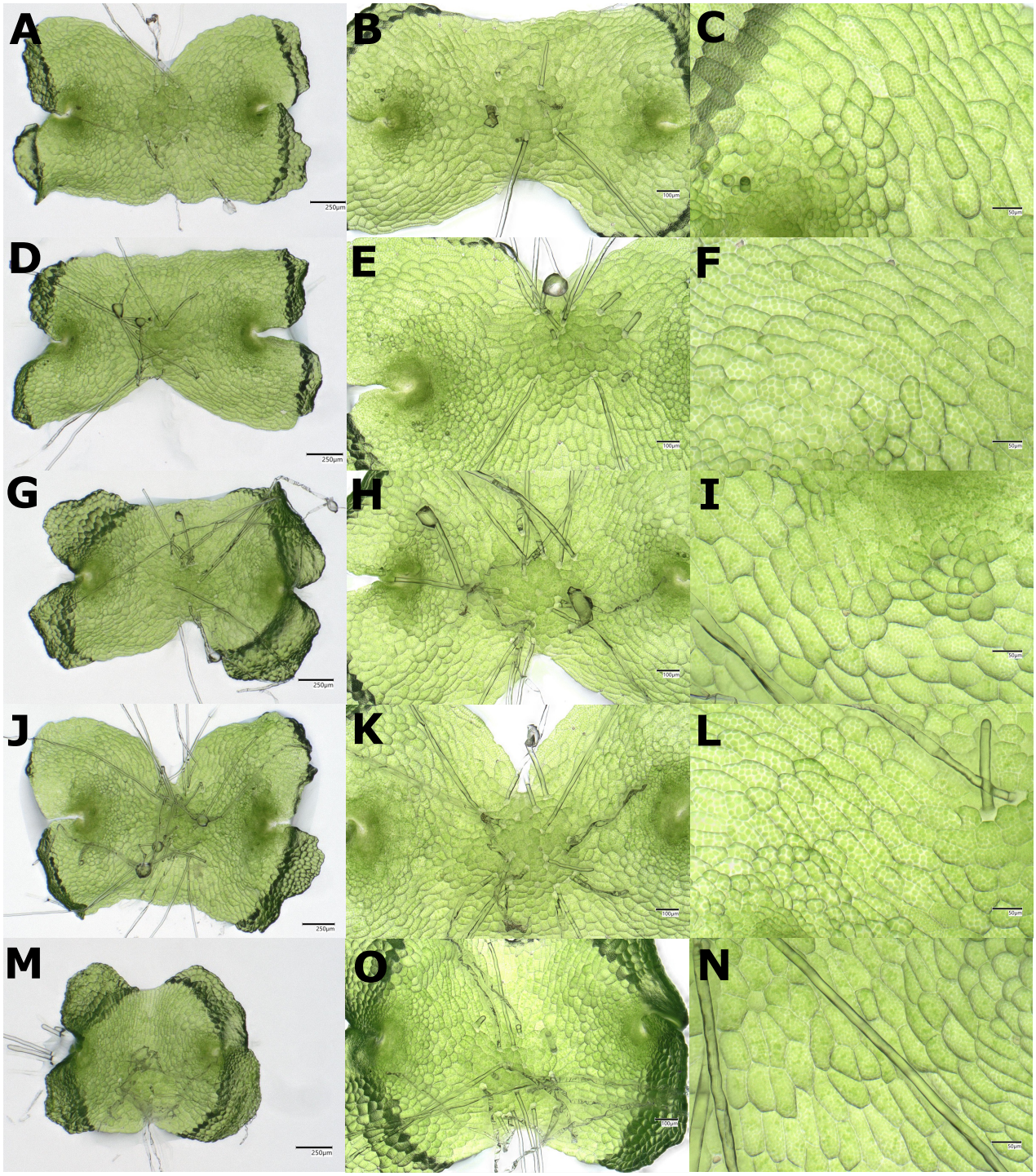


Fig. 4.14. Treatment with peptidoglycan biosynthesis inhibiting antibiotics induces chloroplast and cell defects

M. polymorpha, Tak-1 gemmae were grown for 3 days on media supplemented with peptidoglycan biosynthesis inhibiting antibiotics: (A-C) Control; (D-F) 220 µM Cefotaxime; (G-I) 100 µM Fosfomycin; (J-L) 300 µM Ampicillin; (M-N) 20 µM Vancomycin. Gemmalings were then imaged with a Keyence VHX-7000.

The images of antibiotic-treated gemmalings were used to quantify differences in chloroplast and cell dimensions and compared to untreated plants (Fig. 4.15). Chloroplasts and cells were measured using the image analysis software FIJI. More than 40 chloroplasts and 20 cells were measured for each antibiotic treatment from multiple gemmalings and were sampled from flattest, most mature tissue in each gemmaling. Measurements were compared to those in untreated Tak-1 gemmalings using a one-way ANOVA and Dunnett's multiple comparison test.

Treatment with cefotaxime and ampicillin increased chloroplast size significantly in Tak-1 gemmalings (Fig. 4.14). Both the longest and shortest axes of chloroplasts treated with cefotaxime and ampicillin were larger than those in untreated Tak-1 gemmalings ($p < 0.0001$) (Fig. 4.15 A and B). However, the long and short axes of the chloroplasts in gemmalings treated with fosfomycin and vancomycin did not significantly differ from those in untreated Tak-1 gemmalings ($p > 0.05$). To assess changes in chloroplast shape, we calculated the long axis/short axis ratio for each antibiotic treatment (Fig. 4.15 C). Although the longest and shortest axes of chloroplasts in vancomycin-treated gemmalings remained unchanged, the ratio between these values was significantly higher than in untreated Tak-1 gemmae ($p = 0.0054$), suggesting a slight reduction in chloroplast circularity. The long/short axis ratio was also higher in both cefotaxime- and ampicillin-treated gemmalings compared to untreated Tak-1 plants ($p < 0.005$). Conversely, there was no significant change in this ratio when gemmalings were treated with fosfomycin ($p = 0.3487$), indicating no alteration in chloroplast size or shape with this antibiotic treatment.

Fosfomycin treatment didn't alter chloroplast size or shape but led to fewer chloroplasts per cell ($p = 0.0015$) and reduced cell area ($p = 0.0244$) in gemmalings compared to untreated ones (Fig. 4.15 D and E). However, the number of chloroplasts per unit cell area remained consistent between fosfomycin-treated and untreated gemmalings ($p = 0.5808$) (Fig. 4.15 F), suggesting this change was related to reduced cell size. In contrast, vancomycin treatment had no significant effects on chloroplast or cell morphology ($p > 0.05$). Cefotaxime and ampicillin treatments resulted in fewer chloroplasts per cell compared to untreated Tak-1 gemmalings ($p < 0.0001$), likely due to their larger size. Additionally, cell area was smaller in both cefotaxime- and ampicillin-treated gemmalings compared to untreated ones ($p =$

0.0028 and p = 0.0001, respectively). The reduced chloroplast abundance in these treatments was not solely due to smaller cell size, as the number of chloroplasts per unit cell area was also lower compared to untreated Tak-1 gemmalings ($p < 0.0001$).

To summarise, treatment with different peptidoglycan biosynthesis inhibiting antibiotics had varying effects on *M. polymorpha* plants. Ampicillin and cefotaxime, which have similar mechanisms of action, had the most significant effect, causing defective chloroplasts (differences in all of the chloroplast and cell parameters measured). It is therefore likely that both cefotaxime and ampicillin strongly inhibit the biosynthesis of peptidoglycan in *M. polymorpha*. On the other hand, more subtle, if any, effects on chloroplast and cell measurements were observed in plants treated with fosfomycin and vancomycin. However, in both case, differences compared to untreated *M. polymorpha* plants were observed suggesting some level of peptidoglycan biosynthesis inhibition is occurring.

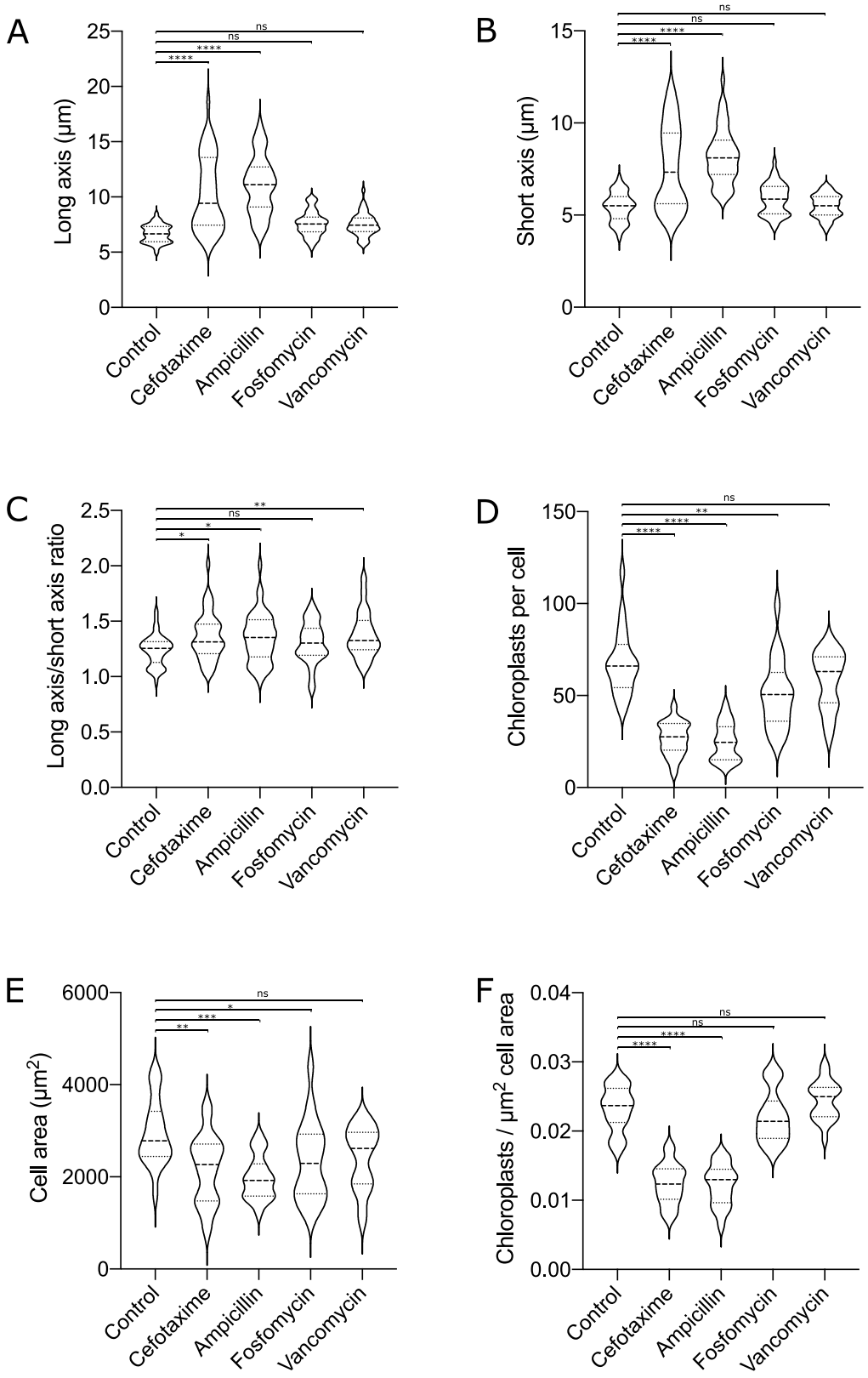


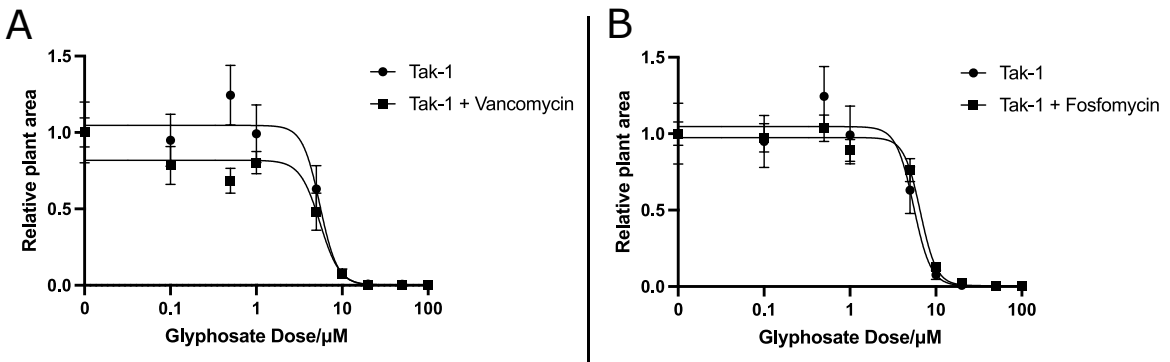
Fig. 4.15. Effect of peptidoglycan biosynthesis inhibiting antibiotics on chloroplast size and abundance and cell size in wild-type *M. polymorpha*

Violin plots of various chloroplast and cell measurements of untreated Tak-1 (Control) and antibiotic-treated Tak-1 gemmalings were calculated using the image analysis software FIJI. The chloroplast longest axis (A), shortest axis (B), long/short axis ratio (C), chloroplast abundance per cell (D), cell area (E) and chloroplast abundance per unit area of cell (F) were measured or calculated. A one-way ANOVA and Dunnett's multiple comparison test was performed to compare values in treated gemmalings to untreated control gemmalings, *=p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001, n.s. = not significant.

4.4.10. Treatment of *M. polymorpha* with peptidoglycan biosynthesis inhibiting antibiotics does not alter sensitivity to glyphosate

M. polymorpha Mp*murF* predicted loss-of-function mutants are not more sensitive to glyphosate than wild-type plants. This is consistent with MurA conferring tolerance to glyphosate in *M. polymorpha* independently of peptidoglycan biosynthesis. To further test this, I co-treated Tak-1 *M. polymorpha* plants with the peptidoglycan biosynthesis inhibiting antibiotics cefotaxime, ampicillin, fosfomycin and vancomycin in conjunction with glyphosate (Fig. 4.16). The effect of these antibiotics on Tak-1 gemmaling phenotype was assessed in the previous section. Tak-1 plants grown on media not treated with an antibiotic were used as the control. All plants were grown for 14 days before imaging and living tissue area calculated. The resulting data was fitted with a four-parameter log-logistic equation and from this, the IC_{50} for each antibiotic co-treatment and control was calculated for glyphosate (Fig. 4.16 A-D).

No significant differences were found between the IC_{50} values for Tak-1 treated with glyphosate alone and Tak-1 co-treated with any of the four antibiotics tested (Welch's t-test, $p > 0.05$) (Fig. 4.16 A-D). This result is consistent with MpMurA, rather than peptidoglycan biosynthesis, being responsible for glyphosate tolerance in *M. polymorpha*. Vancomycin caused reduced plant area at most glyphosate doses compared to the control, consistent with the phenotype observed in Section 4.4.9. However, the glyphosate dose required to inhibit vancomycin-treated plants by the same proportion remained the same (Fig. 4.16 A). Fosfomycin, despite inhibiting MurA, did not exhibit a synergistic effect with glyphosate, possibly due to inefficient inhibition of MpMurA by fosfomycin, as indicated by the lack of a macrochloroplast phenotype in Section 4.4.9. These findings support glyphosate tolerance in *M. polymorpha* being independent of peptidoglycan biosynthesis and suggest that some attribute of the MurA protein contributes to glyphosate tolerance.

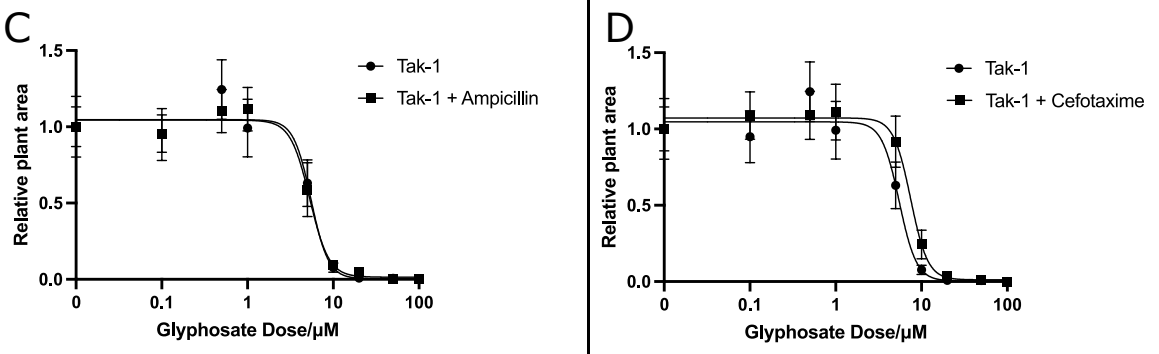


Vancomycin

Treatment	IC_{50} (µM)	P-value
Control	5.459 ± 0.5673	0.7349
20 µM Vancomycin	5.760 ± 0.6822	

Fosfomycin

Treatment	IC_{50} (µM)	P-value
Control	5.459 ± 0.5673	0.1074
100 µM Fosfomycin	6.643 ± 0.4613	



Ampicillin

Treatment	IC_{50} (µM)	P-value
Control	5.459 ± 0.5673	0.8117
300 µM Ampicillin	5.281 ± 0.4845	

Cefotaxime

Treatment	IC_{50} (µM)	P-value
Control	5.459 ± 0.5673	0.0798
220 µM Cefotaxime	7.355 ± 0.9121	

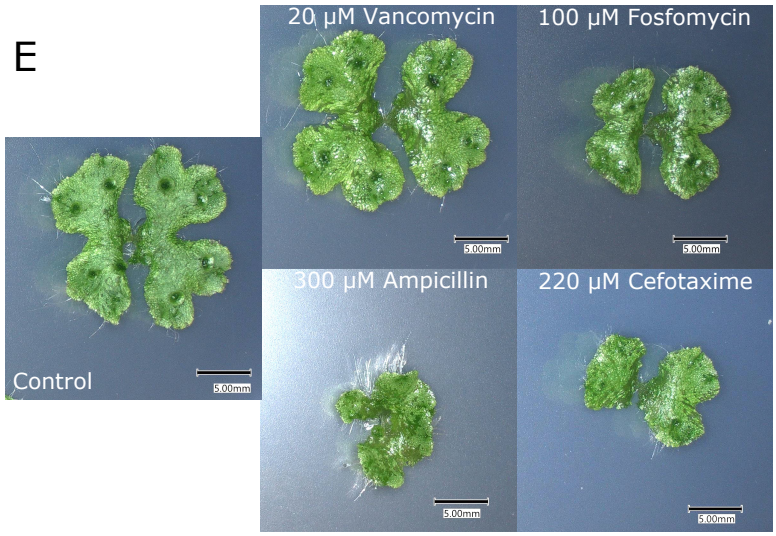


Fig. 4.16. Co-treatment with glyphosate and peptidoglycan biosynthesis inhibiting antibiotics does not increase sensitivity to glyphosate

(A-D) Dose response curves of the plant areas of *M. polymorpha* Tak-1 plants grown for 14 days on plates co-treated with glyphosate and one of the antibiotics cefotaxime, ampicillin fosfomycin or vancomycin. The data are fitted with four-parameter log-logistic regression curves Error bars are ±SE. IC_{50} values were interpolated from the curves and compared to those from plates without co-treatment using a Welch's unpaired t-test.

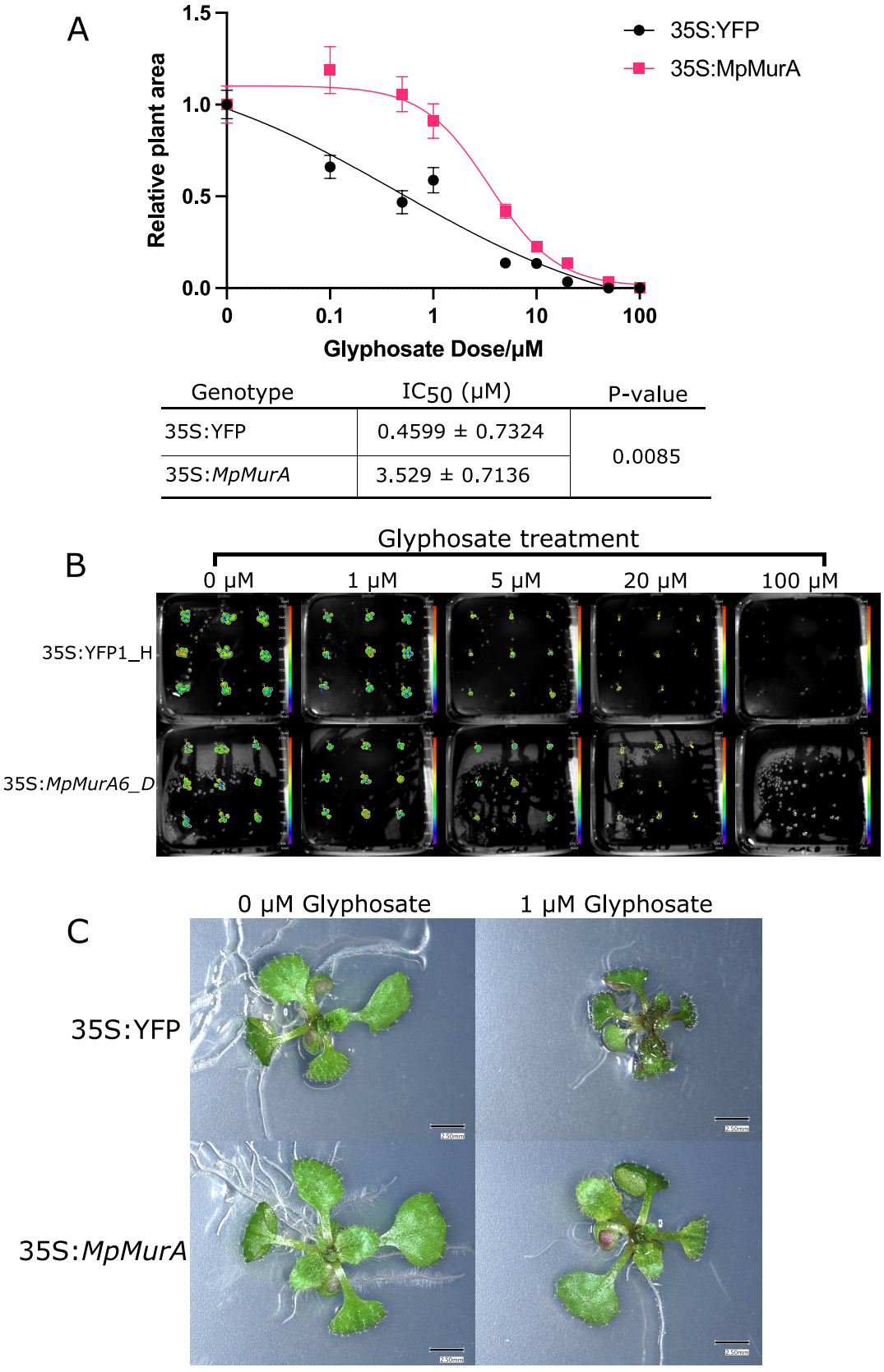
E. Symptoms of Tak-1 plants on plates treated with the four antibiotics tested plus the control taken with a Keyence VHX-7000.

4.4.11. Expression of the Mp*MurA* gene confers glyphosate resistance in *A. thaliana*

To further test if MpMurA contributes to glyphosate tolerance in *M. polymorpha*, MpMurA was expressed in the angiosperm *A. thaliana*. *A. thaliana*, like other angiosperms, does not encode a MurA enzyme. MpMurA was expressed in *A. thaliana* under the CaMV 35S promoter. The YFP protein was also expressed in *A. thaliana* under the same promoter as a control.

The resultant dose response curves are presented in (Fig. 4.17. A). The dose responses to glyphosate were measured for 3 independent *A. thaliana* lines expressing both YFP and MpMurA, and the values pooled. The area of 35S:YFP (control) plants grown on glyphosate was smaller than when grown on no glyphosate conditions. Almost no growth was measured when 35S:YFP were grown on 10 µM glyphosate. These plants were still alive however, determined by autofluorescence measured in leaves. Some autofluorescence was detected up to 20 µM of glyphosate (Fig. 4.17 B). Both 50 and 100 µM of glyphosate was lethal to 35S:YFP plants after 14 days of treatment, suggesting a lethal dose to these lines of between 20 and 50 µM glyphosate. *A. thaliana* plants expressing 35S:Mp*MurA* were, on average, larger than those expressing 35S:YFP across all doses of glyphosate. Furthermore, the growth was not reduced in 35S:Mp*MurA* plants until a dose of between 1 and 5 µM of glyphosate. Some living plant tissue of 35S:Mp*MurA* plants was detected at a glyphosate dose of 50 µM, suggesting a higher lethal dose for 35S:Mp*MurA* than 35S:YFP lines of between 50 and 100 µM of glyphosate. These data are consistent with exogenous expression of Mp*MurA* conferring glyphosate resistance in *A. thaliana*.

Glyphosate IC_{50} values for 35S:YFP and 35S:MpMurA plants were determined using four-parameter log-logistic regressions (Fig. 4.17 A). 35S:MpMurA plants had a 7.7-fold higher IC_{50} compared to 35S:YFP plants. A Welch's t-test confirmed significant glyphosate resistance in 35S:MpMurA plants ($p = 0.0085$). This increased resistance was visually apparent even at low glyphosate doses (Fig. 4.17 C). This clear increase in resistance to glyphosate in *A. thaliana* plants expressing Mp*MurA* is strong evidence for the contribution of MurA to glyphosate tolerance in *M. polymorpha*.



Genotype	IC_{50} (µM)	P-value
35S:YFP	0.4599 ± 0.7324	0.0085
35S:*MpMurA*	3.529 ± 0.7136	

Fig. 4.17. *A. thaliana* plants expressing MpMurA are more resistant to glyphosate
A. Dose response curve of the plant area of *A. thaliana* expressing 35S:YFP and 35S:At*MurA*. Plants were grown for 14 days on glyphosate-treated plates. The data are fitted with a four-parameter log-logistic regression curve. Error bars are ±SE. IC_{50} values of 35S:*MurA* plants are significantly higher than that for a 35S:YFP plants.
B. Images of chlorophyll auto-fluorescing area of *A. thaliana* plants on 5 selected doses of glyphosate.
C. Brightfield images of 35S:YFP and 35S:Mp*MurA* plants on control and 1 µM glyphosate plates taken with a Keyence VHX-7000 microscope.

4.4.12. The overexpression of partial forms of MpMurA in *M. polymorpha* does not increase glyphosate resistance

I have found strong evidence that MpMurA causes glyphosate tolerance in *M. polymorpha* and can confer resistance when expressed in *A. thaliana*. This mechanism of tolerance is independent from peptidoglycan biosynthesis. One way in which MpMurA may cause glyphosate tolerance in *M. polymorpha* is by binding glyphosate, and in doing so, make it less available to inhibit MpEPSPS. To test this, I overexpressed mutated genes coding for various versions of the MpMurA enzyme, the structure of which is predicted in Supplementary Fig. S4.4. These alterations to the expressed MpMurA enzyme were designed to potentially affect the enzymes interaction with glyphosate. Each iteration of the enzyme was designed to be enzymatically inactive, but to still have varying portions of the active site intact (Supplementary Fig. S4.3). This was to test if glyphosate binding alone, and not enzyme activity of MpMurA contributed to glyphosate tolerance in *M. polymorpha*. The construct 35S:Mp*murA*_cut1 was designed to express and MurA protein missing 42 amino acids at the C-terminus of the enzyme. The AlphaFold predicted structure of this protein shows most of the active site of the protein to be intact, but with many structural alterations which likely disrupt the activity of the protein. 35S:Mp*murA*_cut2 encodes a more a more severely truncated version of MpMurA, missing one of the two domains which comprise the MurA enzyme and therefore has very large changes to the active site. However, this protein still encodes the cysteine at position 166 which is essential for substrate binding in MurA enzymes and therefore potentially glyphosate binding. The final construct, Mp*murA*_cut3, is predicted to express a protein with the same truncation at the C-terminus as 35S:Mp*murA*_cut1 but was also designed to be missing the chloroplast transit peptide. Therefore this protein is likely less active than the wild-type and also not localised correctly.

The Mp*murA* altered genes were expressed in *M. polymorpha* under the CaMV 35S promoter. The YFP protein was also expressed under the same promoter as a control. Transformed plants were then grown for 14 days on varying doses of glyphosate before imaging (Fig. 4.18). The plant area relative to control was measured at 10 µM of glyphosate for lines overexpressing the described constructs and tested for significant differences

compared to the relative plant areas for the wild-type accessions Tak-1 and Tak-2 using a Welch's t-test (Fig. 4.18 A).

The relative plant area at 10 µM glyphosate was higher than Tak-2 for all lines expressing mutated versions of the MpMurA protein. However, in five out of the six of these lines the difference was not significant (Welch's t-test, $p > 0.05$). The line 35S:Mp*murA*_cut2_2 was found to have a significantly larger relative area at 10 µM of glyphosate than Tak-2 ($p = 0.00393$). This difference can clearly been seen in images in the living tissue of the plants taken at this dose, and more 35S:Mp*murA*_cut2_2 plants survived this dose than Tak-2 plants (Fig. 4.18 B). This evidence is consistent with the 35S:Mp*murA*_cut2_2 line being more resistant to glyphosate than Tak-2. On the other hand, none of the lines overexpressing altered forms of MpMurA were more resistant to glyphosate than Tak-1 or from the 35S:YFP line included in this experiment ($p > 0.05$). This contrasts the result in Section 4.4.4 which found that overexpressing the full form of MpMurA in *M. polymorpha* led to an increase in glyphosate resistance in many instances compared to Tak-1, Tak-2 and 35S:YFP control lines. This evidence suggests that expression of these shortened forms of MpMurA cannot cause glyphosate resistance. Thus is consistent with either an inability to bind glyphosate in these predicted forms, or that the mechanism of glyphosate tolerance utilising MpMurA in *M. polymorpha* requiring the complete MpMurA enzyme.

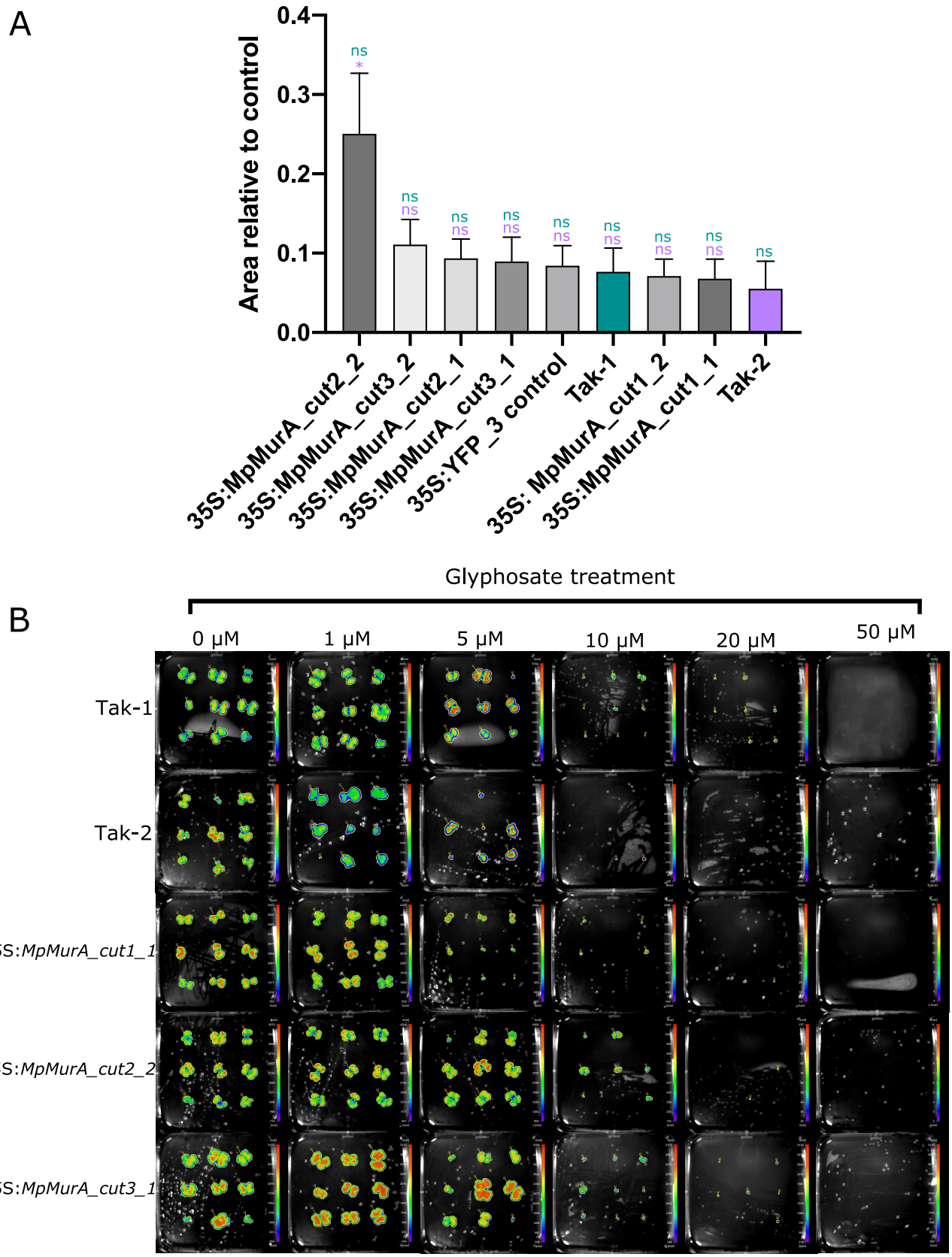


Fig. 4.18. Overexpression of incomplete versions of MpMurA does not increase glyphosate resistance in *M. polymorpha*

A. *M. polymorpha* wild-type and 35S:Mp*MurA*_cut overexpression lines were grown on 0 µM and 10 µM of glyphosate. The auto-fluorescing areas were measured and the ratio between treated and untreated plants calculated and plotted. Error bars are ±SE. Welch's t-tests were performed using Tak-1 and Tak-2 areas relative to control as the baseline *=p < 0.05 and n.s. = not significant.

B. Images of the chlorophyll auto-fluorescing area of Tak-1, Tak-2 and three 35S:Mp*MurA*_cut lines on selected doses of glyphosate.

4.4.13. Micro-scale thermophoresis/Temperature -related intensity change analysis cannot detect glyphosate binding to MpMurA

A hypothesis to explain how MpMurA confers glyphosate tolerance in *M. polymorpha* is that MpMurA binds glyphosate, and in doing so, prevents glyphosate from binding to MpEPSPS and therefore inhibiting the biosynthesis of the aromatic amino acids. To test whether glyphosate does bind to MpMurA, micro-scale thermophoresis/temperature-related intensity change (MST/TRIC) was employed on purified MpMurA and MpEPSPS in the presence of glyphosate (Fig. 4.19) along with the ligands UDP-GlcNAc, S3P and PEP (Supplementary Fig. S4.5).

My experiment in Chapter 3 Section 4.4 shows that glyphosate inhibits purified MpEPSPS enzyme and therefore binds to the enzyme. To test if this binding could be measured with MST/TRIC, labelled EPSPS was titrated with glyphosate in the presence of 1 mM S3P – since the formation of an EPSPS-S3P complex is necessary for glyphosate binding (Kishore and Shah, 1988). A significant MST/TRIC signal with an amplitude of approximately 3 ‰ normalised fluorescence was observed (Supplementary Fig. S4.5 A). The MST trace was evaluated at a hot time of between 0.5 seconds and 1.5 seconds after turning on the infra-red (IR) laser. The resulting binding curve (Supplementary Fig. S4.5 B) could be analysed using PALMIST to give a K_D value of MpEPSPS for glyphosate of approximately 20 µM (68.3 % confidence interval, 8 µM – 44 µM).

MST measurements for MpMurA alone with glyphosate did not show any significant changes in amplitude of normalised fluorescence (Supplementary Fig. S4.5 C). The datapoints were all within a 1 ‰ range, which corresponds to the noise usually observed during experiments. A comparison of the change in normal fluorescence compared between MpEPSPS, which binds glyphosate, and MpMurA is shown in Fig. 4.19. Since there was no significant change in the normalised fluorescence of MpMurA titrated with glyphosate, no prediction of the K_D value of MpMurA for glyphosate could be made, and this evidence is consistent with no binding occurring at all. This contrasts the strong glyphosate binding that was measured to MpEPSPS. It is, however, formally possible that MpMurA must form a complex with another molecule if it is to bind glyphosate. Therefore MST measurements were made for MpMurA

with in the presence of UDP-GlcNAc, S3P and PEP. However, a significant MST signal was observed with none of these potential MpMurA interacting molecules when assayed with MpMurA (Supplementary Fig. S4.5 D). Despite this observation, binding of the molecules directly to MpMurA cannot be excluded, as the binding could occur, but with insufficient signal observed with this method. However, these data are consistent with none of these molecules, even the natural substrates UDP-GlcNAc and PEP binding to purified MpMurA *in vitro*. This data does not support the hypothesis that glyphosate binds to MpMurA, but this cannot be fully excluded as no binding was measured for molecules known to bind to MpMurA.

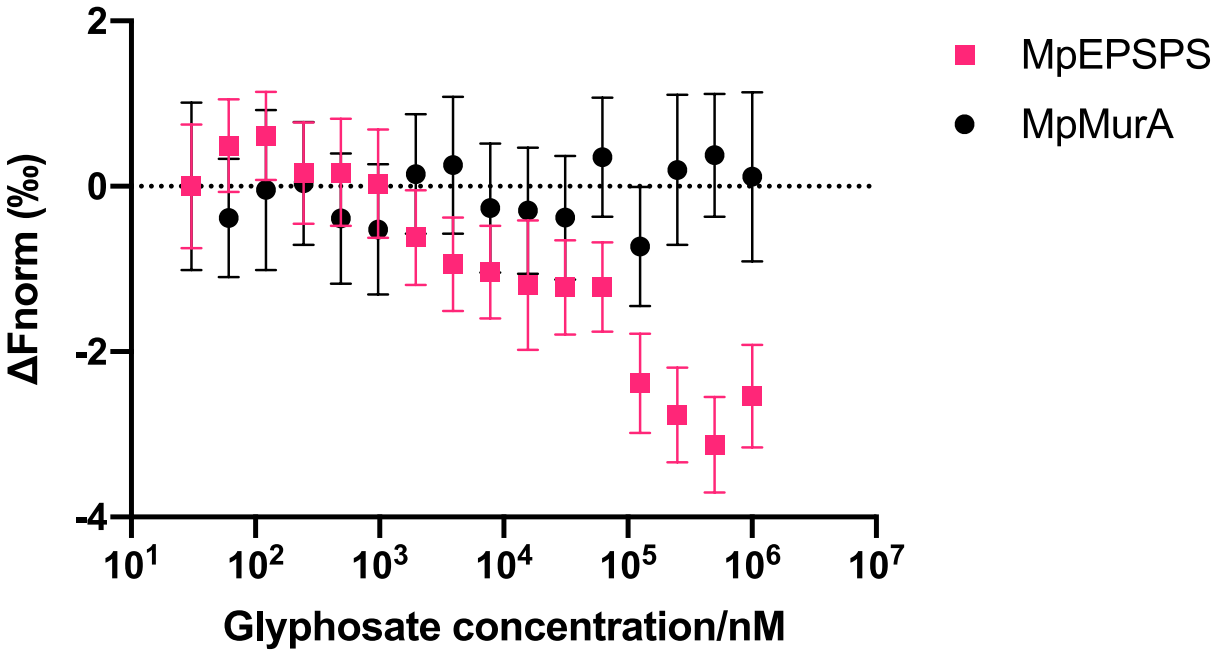


Fig. 4.19. Glyphosate is not found to bind to MpMurA using micro-scale thermophoresis
Change in normalised fluorescence of MpEPSPS (in the presence of 1 mM S3P) and MpMurA with glyphosate at various concentrations up to 1 mM. A clear signal was observed for MpEPSPS and from this a K_D value of approximately 20 µM glyphosate could be calculated. No clear signal was observed for MpMurA with glyphosate which is consistent with glyphosate not binding to MpMurA.

4.4.14. MpMurA catalyses the reaction between shikimate-3-phosphate and phosphoenolpyruvate to produce 5-enolypyruvylshikimate-3-phosphate

MpEPSPS and MpMurA are similar in structure and so it is formally possible that MpMurA is able to catalyse the same reaction as MpEPSPS. This would have the effect of supplementing EPSPS activity, contributing to flux through the shikimate pathway in the presence of glyphosate, and therefore conferring glyphosate tolerance. To test if MpMurA catalyses the reaction of MpEPSPS – the conversion of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) into 5-enolypyruvylshikimate-3-phosphate (EPSP) and phosphate – I measured if there was enzymatic activity from MpMurA with S3P and PEP as substrates (Fig. 4.20). These experiments were conducted with purified MpMurA protein.

I first quantified MpMurA activity by measuring the amount of inorganic phosphate produced in the forward reaction between S3P and PEP. As controls, the activity of purified MpEPSPS for the same reaction was measured, as well as that of MpMurA with its substrates UDP-GlcNAc and PEP. This reaction also leads to the production of inorganic phosphate (Fig. 20 A). The enzyme activity is expressed as µmol phosphate produced/min of reaction time/mg of enzyme (U/mg). At the fixed concentrations of substrates used, activity was detected for EPSPS, and MurA with both the substrates of the EPSPS reaction and the MurA reaction. Therefore, phosphate was produced by MpMurA in the presence of S3P and PEP, which is consistent with MpMurA being able to catalyse the same reaction as EPSPS. However, the rate of reaction of MpMurA was approximately 100-fold less than that of EPSPS. This rate of reaction was approximately 8-fold higher than that of MpMurA with its usual substrates of UDP-GlcNAc and PEP. No accumulation of phosphate was measured when both sets of substrates were incubated without enzyme. Comparisons between the rate of activity of MpEPSPS and MurA, and between MurA with both sets of substrates are difficult to make due to different concentrations of substrates used and different proportions of phosphate produced relative to the other product of the reaction. However, production of phosphate was detected for MpMurA with the substrates of the EPSPS reaction, suggesting the structural similarity between MpMurA and MpEPSPS relates to a shared catalytic activity (Fig. 4.20 A). This suggests that MpMurA is able to transfer the enolpyruvyl moiety of

phosphoenolpyruvate (PEP) to both the 3'-hydroxyl of UDP-GlcNAc and to S3P to create EPSP.

To test if MpMurA catalysed the full reaction between S3P and PEP, and therefore produced EPSP, liquid chromatography-tandem mass spectrometry (LC-MS/MS) was employed to detect for the accumulation of this product (Fig. 4.20 B). EPSP accumulates when MpMurA was incubated with S3P and PEP, consistent with MpMurA catalysing the production of EPSP and therefore sharing catalytic function with EPSPS. However, the quantity of EPSP produced per mg of enzyme, measured in ion counts, was approximately 5500-fold less for MpMurA than measured for MpEPSPS. This suggests that although MpMurA catalyses the reaction of S3P and PEP to form EPSPS, it is much less active than MpEPSPS.

If MpMurA were to contribute to glyphosate tolerance through its production of EPSP from S3P and PEP substrates, tolerance would be increased if its catalysis of this reaction were tolerant to glyphosate. My MST analysis suggests that glyphosate does not bind to MpMurA, while it does bind to MpEPSPS. To test if MurA was inhibited by glyphosate, the amount of EPSP produced by both MpEPSPS and MpMurA was measured at 8 concentrations of glyphosate using (LC-MS/MS) (Fig. 4.20 C). Consistent with my findings in Chapter 3 Section 4.5 the quantity of EPSP produced by MpEPSPS was reduced at increasing concentrations of glyphosate, suggesting inhibition of the forward reaction by the herbicide. However, the quantity of EPSP produced by MpMurA was also reduced by glyphosate, following a similar dose-response relationship as MpEPSPS. A four-parameter log-logistic equation was fitted to the data and from this the IC_{50} for glyphosate of each enzyme interpolated. MpMurA was found to be more sensitive to glyphosate with an IC_{50} of 8.392 µM, whereas MpEPSPS had an IC_{50} of 13.23 µM. These IC_{50} values were found to be statistically significantly different using a student's t-test with a Welch's correction ($p = 0.0445$). Therefore, although MpMurA is active in the reaction of EPSPS, its rate of reaction is much lower, and it is also more sensitive to inhibition with glyphosate than MpEPSPS.

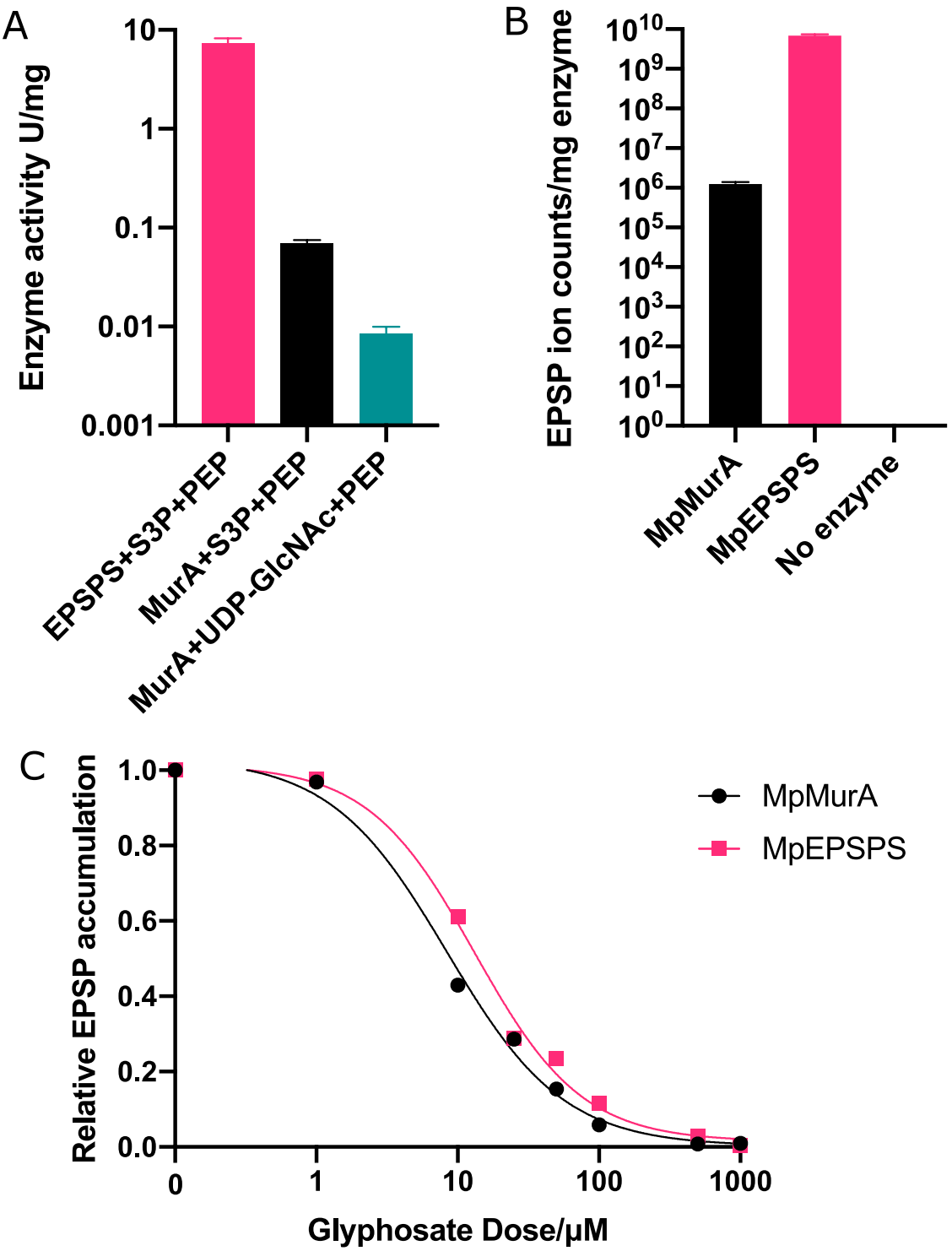


Fig. 4.20. MpMurA catalyses the same reaction as MpEPSPS and is inhibited by glyphosate
A. Enzymatic activity of MpEPSPS and MpMurA with the substrates S3P and PEP and of MpMurA with UDP-GlcNAc and PEP measured as µmol phosphate produced/min of reaction time/mg of enzyme (U/mg).
B. Accumulation of EPSP measured with LC-MS/MS accumulated when S3P and PEP were incubated with MpEPSPS, MpMurA and no enzyme.
C. Dose response curve of the relative accumulation of EPSP measured by LC-MS/MS when S3P and PEP were incubated with MpEPSPS and MpMurA along with varying concentrations of glyphosate. The data are fitted with a four-parameter log-logistic regression curve. The IC_{50} of MpMurA for glyphosate was significantly lower than that of MpEPSPS ($p = 0.0445$).

4.4. Discussion

In this chapter, I aimed to identify the mechanism of glyphosate tolerance in *Marchantia polymorpha*, predicting this to be a form of non-target site resistance. I have identified that the enzyme UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) contributes to glyphosate tolerance in *M. polymorpha*. I show that MpMurA is able to catalyse the conversion of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) into 5-enolypyruvylshikimate-3-phosphate (EPSP) and phosphate. This is the reaction which the target of glyphosate, EPSPS, catalyses. This therefore represents a novel form of tolerance to glyphosate which is likely unique to non-angiosperm streptophytes.

Using BLASTP to query the proteome of *M. polymorpha* with the *A. thaliana* EPSPS amino acid sequence, MpMurA was identified as a significant hit. MurA catalyses the first committed step of peptidoglycan biosynthesis. I produced structural predictions of MpEPSPS and MpMurA using AlphaFold and found the structure of MpMurA to be strikingly similar to that of MpEPSPS. This structural similarity between MpMurA and MpEPSPS likely reflects their functional similarity, with both proteins transferring the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to their respective substrates, uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) and shikimate-3-phosphate (S3P). Glyphosate inhibits EPSPS by binding to the active site in the place of PEP (Schönbrunn et al., 2001). Given the structural likeness between MpMurA and MpEPSPS, and the shared substrate PEP, it is formally possible that glyphosate can also bind in the place of PEP to the MurA enzyme in *M. polymorpha*. This would have the effect of reducing the concentration of EPSPS enzymes inhibited by glyphosate, thus conferring tolerance. The structural similarity of MpMurA to MpEPSPS therefore led me to identify this enzyme as a candidate for conferring glyphosate tolerance in *M. polymorpha*.

Glyphosate is generally used to control angiosperm weeds and few, if any, are naturally tolerant to this herbicide (Duke and Powles, 2008). I hypothesised that MurA contributes to glyphosate tolerance in *M. polymorpha*. If this is the case, I would expect that a complete version of MurA is not present in angiosperms, as they are glyphosate sensitive. To test this hypothesis, I searched the predicted proteomes of 53 species of the Archaeplastida for a

MurA sequence using BLASTP. I also searched the predicted proteomes for an EPSPS sequence. At least one EPSPS enzyme was encoded by the genomes of all the Archaeplastida species queried, except *Azolla filiculloides*. An incomplete version of an EPSPS sequence was retrieved from *A. filliculoides* and this likely represents an artifact of poor proteome coverage of this species rather than it lacking a full version of EPSPS. EPSPS is therefore highly conserved across the Archaeplastida, reflecting its essential role in the shikimate pathway in plants. A full MurA sequence was not identified in the proteomes of any of the angiosperm species I queried. This is consistent with my hypothesis that MurA is involved in glyphosate tolerance in *M. polymorpha*. A partial version of MurA was identified in the proteome of *Solanum lycopersicum*, however, this had a large sequence dissimilarity from other land plant MurA enzymes identified (Supplementary Fig. S4.1). MurA was highly conserved among the non-angiosperm streptophytes, with partial versions or absences in just 3 of the 33 non-angiosperm streptophytes queried. This suggests MurA is highly conserved among the non-angiosperm streptophytes and therefore I predict that these species will also be glyphosate tolerant.

I conducted a glyphosate dose response experiment, involving multiple plant species, in order to test if the presence of a MurA enzyme correlated with glyphosate tolerance. I tested the sensitivity of glyphosate in the glyphosate sensitive angiosperm *A. thaliana*, which lacks a MurA enzyme, the glyphosate tolerant *M. polymorpha*, which encodes a MurA enzyme, and two other species of non-angiosperm streptophytes which encode a MurA enzyme but the sensitivity to glyphosate of which is unknown – *Physcomitrium patens* and *Lunularia cruciata*. Both *P. patens* and *L. cruciata* were more tolerant to glyphosate than *M. polymorpha*, showing little to no inhibition of growth following treatment with concentrations of up to 100 µM of glyphosate. A concentration of less than 10 µM glyphosate was lethal to *A. thaliana*, but 100 µM of glyphosate was not lethal to any of the species encoding a MurA enzyme, despite the significant inhibition of growth observed in *M. polymorpha*. This observation is consistent with the hypothesis that MurA contributes to glyphosate tolerance in *M. polymorpha*, as all plants encoding a MurA enzyme were more toelrant to glyphosate than *A. thaliana*, a species lacking this enzyme. However, the evolutionary relatedness of the three bryophyte species raises the possibility of other shared adaptations contributing to their common glyphosate tolerance.

I produced ten transgenic *M. polymorpha* lines overexpressing the Mp*MurA* gene. This resulted in an increase in the glyphosate resistance in these lines which provides strong evidence to support the role of MurA in conferring glyphosate tolerance in *M. polymorpha*. The majority of *M. polymorpha* plants overexpressing Mp*MurA* showed less growth inhibition at a concentration of 10 µM glyphosate compared to Tak-1, Tak-2 and control plants expressing YFP. Among the lines overexpressing Mp*MurA*, the relative area of five of the lines was significantly greater than Tak-1, the most glyphosate tolerant of the wild-type lines. Additionally, the mean relative areas were greater in three other lines, although statistical significance was not reached. This suggests overexpression of Mp*MurA* increases glyphosate resistance. Taken together, this evidence strongly supports the hypothesis that MurA contributes to glyphosate tolerance in wild-type *M. polymorpha* plants.

Using CRISPR-Cas9 mutagenesis, I generated seven Mp*murA* predicted loss-of-function mutants. Five of these mutants were predicted to be weak loss-of-function, as the mutations were predicted to cause the expression of a MurA enzyme with few changes to its amino acid sequence. Two mutants were predicted to be MpMurA strong loss-of-function, as they had mutations which resulted in a frame shift and a large truncation of the expressed protein due to a premature stop codon. In *P. patens,* loss-of-function mutations in its *MurA* genes leads to a macrochloroplast phenotype (Homi et al., 2009; Machida et al., 2006). I measured the dimensions and abundance of the chloroplasts in three of my Mp*murA* predicted loss-of-function mutants to test for macrochloroplasty. I found large increases in the size of the chloroplasts and many times fewer chloroplasts in my Mp*murA* mutants. This is strong evidence of a loss-of-function of MpMurA in these mutants which leads to disruption in chloroplast division and macrochloroplasts as a result of the disruption of peptidoglycan biosynthesis. There was a large increase in chloroplast size in both the predicted weak and strong loss-of-function mutants, suggesting that even a small change to the amino acid sequence of MpMurA leads to large disruption of this enzyme's activity. Therefore I predicted that the mutations in the Mp*MurA* gene in the seven mutant lines all led to a severe loss-of-function of MpMurA.

All the Mp*murA* predicted loss-of-function lines were found to be more sensitive to glyphosate than both Tak-1 and Tak-2 wild-type lines. Tak-2 is more sensitive to glyphosate than Tak-1, and therefore Tak-2 was taken as conservative baseline for statistical comparisons to be made. All Mp*murA* predicted loss-of-function mutant lines had a significantly smaller relative area than Tak-2 plants. Furthermore, there was no correlation between the magnitude of the predicted mutation in the Mp*MurA* gene and the increased sensitivity to glyphosate observed. This may be a result of all mutations generated resulting in a complete loss-of-function, and the variance observed being due to genetic diversity in the plants due to the cross between Tak-1 and Tak-2 which produced spores used for transformation to generate the Mp*murA* loss-of-function lines. Therefore, mutations leading to a loss-of-function of MpMurA cause *M. polymorpha* plants to be more sensitive to glyphosate than both Tak-1 and Tak-2 wild-type plants. This is strong evidence to support the role of MurA in glyphosate tolerance in wild-type *M. polymorpha*.

MurA catalyses the first committed step of peptidoglycan biosynthesis. It is formally possible that changes in rates of peptidoglycan biosynthesis in MpMurA overexpressing and Mp*murA* loss-of-function lines resulted in the observed changes in glyphosate sensitivity. To determine if MpMurA contributes to glyphosate tolerance in *M. polymorpha* independently of peptidoglycan biosynthesis, I generated five Mp*murF* loss-of-function mutants using CRISPR-Cas9 mutagenesis. Each of these mutants was predicted to express a MurF protein with a large truncation which I predicted would lead to a disruption of peptidoglycan biosynthesis. Consistent with this, Mp*murF* mutants exhibited a macrochloroplast phenotype and there were many times fewer chloroplasts in the cells of Mp*murF* mutants compared to wild-type. This suggests that either MpMurA and MpMurF are both essential for chloroplast division, or that peptidoglycan is in some way required for chloroplast division. Furthermore, macrochloroplasty may not be the result of disrupted chloroplast division, but instead due to structural changes to the chloroplast wall due to a lack of peptidoglycan. The severe phenotype observed in Mp*murF* mutants is consistent with a large or complete loss-of-function of the MurF enzyme.

I found that Mp*murF* loss-of-function mutants were mostly not significantly more or less sensitive to a concentration of 5 µM glyphosate than the wild-type lines Tak-1 and Tak-2.

Only one statistically significant change in relative plant area was identified, in which *murF_90crsc_12* was less inhibited by glyphosate compared to Tak-2. The slightly larger relative area of this line compared to Tak-2 suggests it is slightly more resistant to glyphosate. However, this line was not more resistant to glyphosate than Tak-1 and so the difference in resistance observed between *murF_90crsc_12* and Tak-2 may not be due to the loss-of-function of MurF. Furthermore, when a dose-response assay was conducted with *murF_90crsc_12*, it showed a similar dose response relationship to glyphosate as both Tak-1 and Tak-2, suggesting the magnitude of resistance in this line is similar to that of wild-type plants. Taken together, this is strong evidence that a disruption in peptidoglycan biosynthesis and resultant macrochloroplasty are insufficient to influence glyphosate sensitivity in *M. polymorpha*. This evidence is therefore consistent with MurA acting independently of peptidoglycan biosynthesis in *M. polymorpha* to confer glyphosate tolerance.

Treatment of non-angiosperm plants with peptidoglycan biosynthesis inhibiting antibiotics, active against certain bacteria, can lead to macrochloroplasty and disrupted chloroplast division (Hirano et al., 2016; Katayama et al., 2003; Matsumoto et al., 2012; Tounou et al., 2002). These antibiotics therefore likely inhibit peptidoglycan biosynthesis in these plant species. I found that treatment of *M. polymorpha* gemmae with cefotaxime and ampicillin, led to macrochloroplasts and fewer chloroplasts, consistent with these antibiotics inhibiting peptidoglycan biosynthesis in *M. polymorpha*. Cefotaxime and ampicillin inhibit peptidoglycan biosynthesis in bacteria by binding to and inhibiting penicillin-binding proteins (PBPs). *M. polymorpha* gemmae were also treated with fosfomycin which inhibits MurA and vancomycin which prevents certain peptides from being incorporated into the peptidoglycan matrix. No change to the size of the chloroplasts was measured in plants treated with these antibiotics, but fosfomycin treatment caused a reduction in cell size and chloroplast abundance while vancomycin caused a change in chloroplast shape. These antibiotics, at the dose used, appear to have a less pronounced inhibition of peptidoglycan biosynthesis in *M. polymorpha*, though the subtle changes observed likely reflect some inhibition of peptidoglycan biosynthesis. Furthermore, after two weeks grown on the selected antibiotics, all *M. polymorpha* plants had a similar symptom of a glossy thallus surface. This may be a result of a change in cuticle composition as a result of the inhibition of peptidoglycan biosynthesis. These data are consistent with all of the antibiotics effecting *M. polymorpha* in

some way, with ampicillin and cefotaxime causing the macrochloroplasty which is linked to inhibition of peptidoglycan biosynthesis in plants. However, when *M. polymorpha* plants were co-treated with these peptidoglycan biosynthesis inhibiting antibiotics and glyphosate, no difference in the toxic effect compared glyphosate alone was observed. This is further evidence consistent with MurA acting independently of peptidoglycan biosynthesis to confer glyphosate tolerance in *M. polymorpha*.

A MurA enzyme is not encoded in the genome of *A. thaliana* and therefore this species could be used to test if MurA does indeed confer glyphosate tolerance in *M. polymorpha* independently from peptidoglycan biosynthesis. I found that exogenous overexpression of Mp*MurA* in *A. thaliana* led to an increase in glyphosate resistance in this species. Overexpression of Mp*MurA* in *A. thaliana* led to a more than 7-fold increase of the IC_{50} to glyphosate compared to *A. thaliana* plants expressing YFP instead of MurA. Furthermore, the lethal dose of glyphosate to *A. thaliana* plants expressing MurA was higher. Taken together, these data are consistent with MurA conferring glyphosate-tolerance independently of peptidoglycan biosynthesis. They are also consistent with the conclusion that MurA confers glyphosate tolerance in wild-type *M. polymorpha*.

I have shown strong evidence supporting the role of MpMurA in glyphosate tolerance, both in *M. polymorpha* and the heterologous system of *A. thaliana*. I have also shown that this mechanism of tolerance likely acts independently of peptidoglycan biosynthesis. To test if MpMurA activity or glyphosate binding to the enzyme, caused glyphosate tolerance, I overexpressed various forms of MpMurA in *M. polymorpha* designed to be enzymatically inactive, yet retain varying portions of the active site to which glyphosate may bind. There was a small increase in the level of resistance to glyphosate of all seven of the lines overexpressing MurA variants compared to Tak-2, and an increase in resistance of five of the lines compared Tak-1. However, this increase in resistance was only significant in one of the lines and only when compared to the more glyphosate sensitive wild-type line Tak-2. This line (35S:Mp*murA*_cut2_2) expressed the most severely altered MpMurA enzyme which was missing an entire domain of the protein. However, this left exposed the cysteine 166 (115 when numbered according to the *Escherichia coli* MurA) which has been proposed to be essential for binding PEP and the antibiotic fosfomycin, which is structurally similar to

glyphosate (Eschenburg et al., 2005; Krekel et al., 1999; Skarzynski et al., 1996). This alteration to MpMurA may allow glyphosate to bind to this residue in MurA, and in doing so make it less available to bind to and inhibit EPSPS, conferring tolerance. However, this is inconsistent with the observation of the other line expressing the same version of MurA which is not more resistant. Furthermore, the slight increase in resistance observed in the 35S:Mp*murA*_cut2_2 line, and the other lines expressing variants of MpMurA, is not as large as the increase in resistance observed following overexpression of the full MpMurA enzyme. This is consistent with mechanisms of resistance involving MurA in *M. polymorpha* requiring the full MpMurA enzyme and therefore potentially the activity of MpMurA.

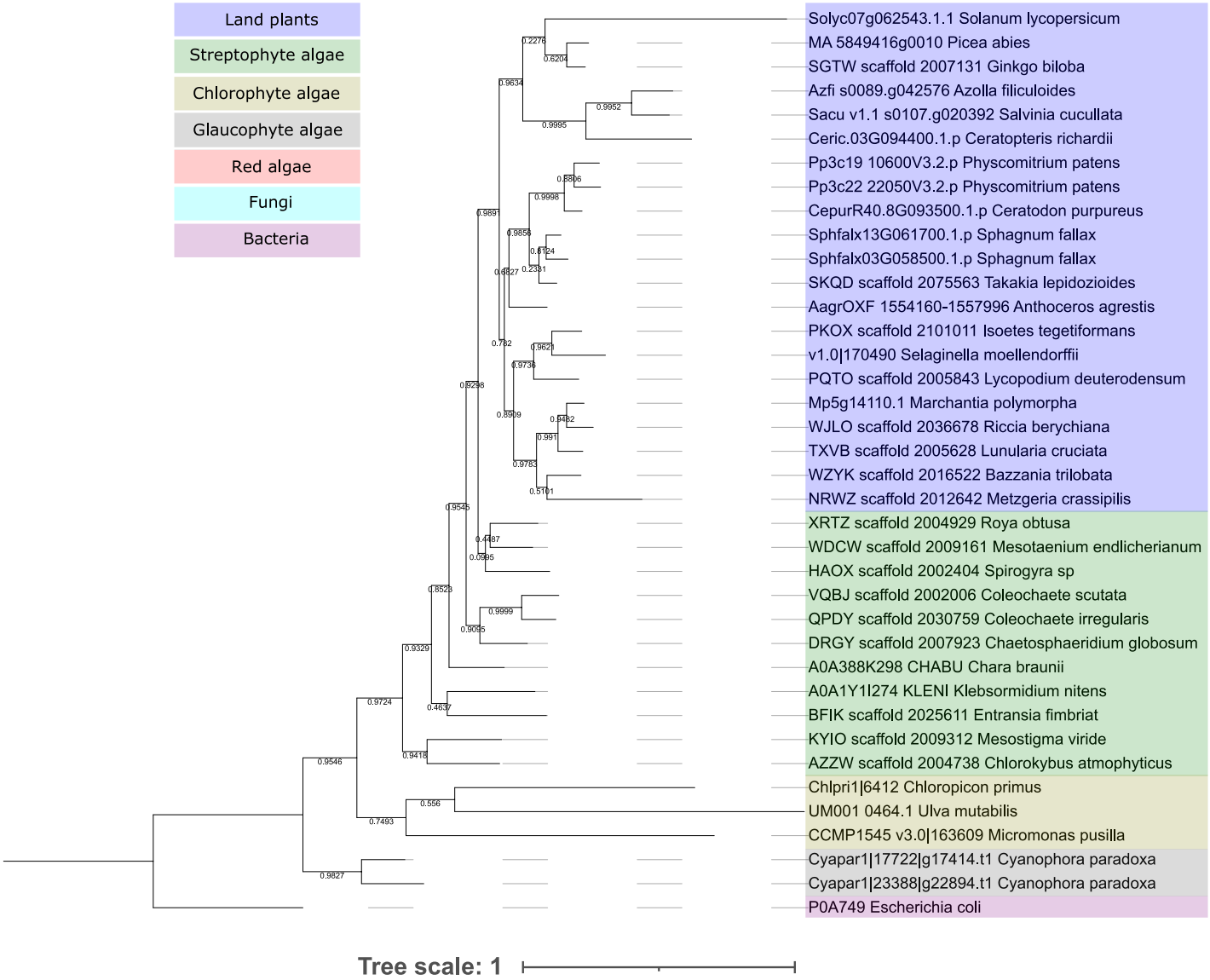
No binding of glyphosate to purified MpMurA was observed using micro-scale thermophoresis/temperature-related intensity change (MST/TRIC). This contrasts with the strong binding of glyphosate to MpEPSPS observed using the same method. This is inconsistent with the hypothesis that MpMurA may bind to glyphosate, and in doing so make it less available to go on and bind to MpEPSPS, therefore conferring glyphosate tolerance. However, no evidence of binding was also observed between MpMurA and its substrates UDP-GlcNAc and PEP. These substrates are known to bind to MurA and therefore the lack of binding to these substrates and potentially glyphosate may reflect a limitation of this technique (López-Méndez et al., 2021b, 2021a; Scheuermann et al., 2016). In summary, evidence from MST measurements were not consistent with glyphosate binding to MurA, but potential false negatives produced by the technique mean that this still cannot be ruled out.

I have shown evidence that supports the role of an active version of MpMurA being required to confer glyphosate tolerance in *M. polymorpha*. Given the structural similarity between MpMurA and MpEPSPS, I hypothesised that MpMurA may have the capacity to catalyse the same reaction as MpEPSPS, thereby supplementing EPSPS activity and facilitating flux through the shikimate pathway even in the presence of glyphosate. By independently measuring both of the products of the EPSPS reaction I showed that MpMurA catalyses the conversion of S3P and PEP into EPSP and phosphate. This is consistent with the MpMurA being able to transfer the enolpyruvyl moiety of PEP to both UDP-GlcNAc and S3P. In fact, MurA appears to be catalytically more active with S3P as a substrate than UDP-GlcNAc.

However, the catalytic activity of MpMurA was substantially lower than that of MpEPSPS. Furthermore, glyphosate inhibits MpMurA, with the rate of EPSP produced in the forward reaction showing a dose-dependent relationship with glyphosate. MpMurA was more sensitive to glyphosate than MpEPSPS and this contradicts previous reports that no other enzyme than EPSPS is inhibited by glyphosate (Pollegioni et al., 2011). However, MpMurA sensitivity to glyphosate is consistent with glyphosate binding to MpMurA. This likely reduces the availability of glyphosate to bind to MpEPSPS – conferring tolerance. MpMurA activity with S3P and PEP as substrates may also contribute to glyphosate tolerance in *M. polymorpha*, as if the expression level of MpMurA is sufficient, this may increase flux through the shikimate pathway enough that it confers tolerance to glyphosate. This would represent an analogous mechanism to target site overexpression-based resistance which has evolved in many weed species (reviewed in Gaines et al., 2020).

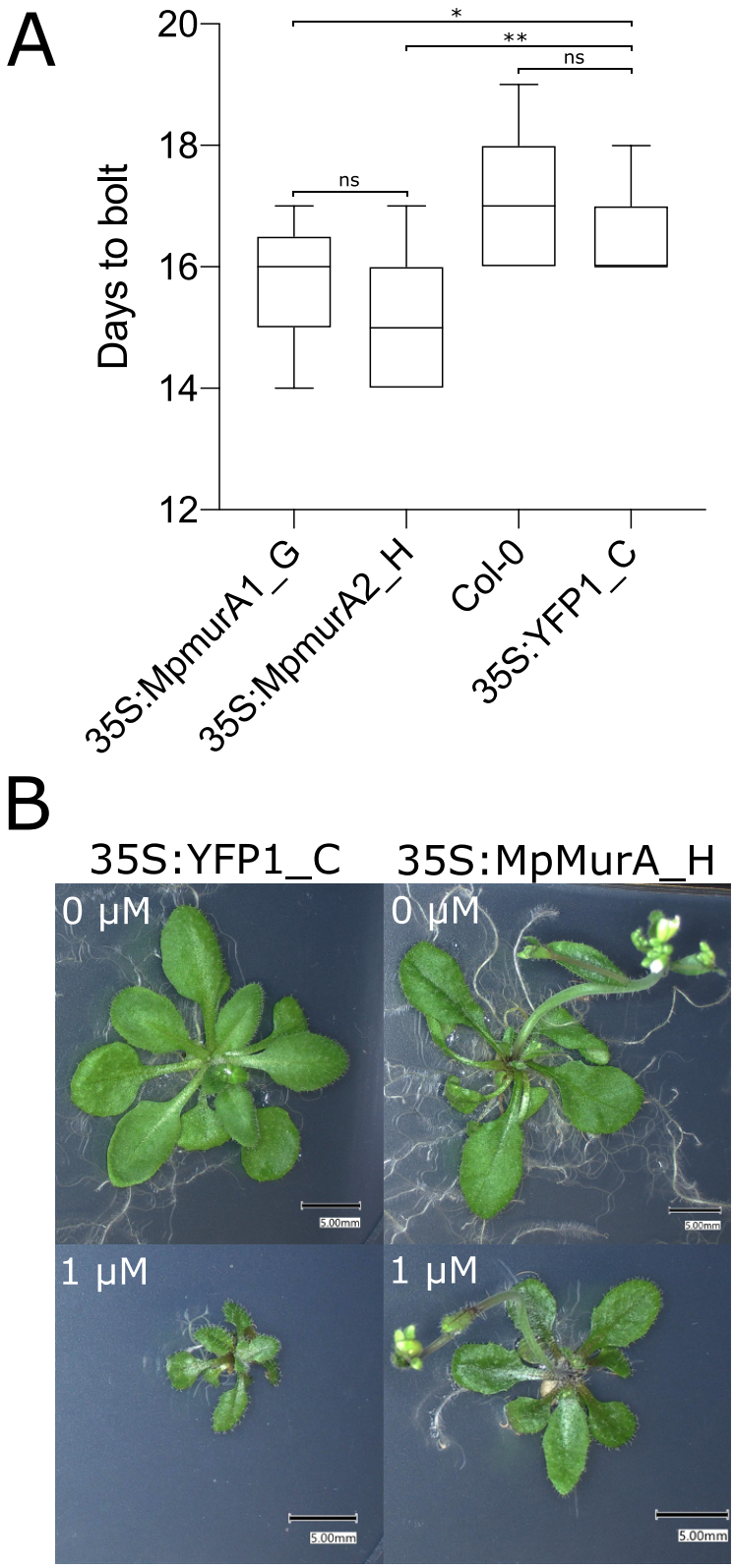
In conclusion, I identified MurA as a candidate for conferring glyphosate tolerance in *M. polymorpha*. I have shown that Mp*MurA* overexpression lines are more resistant to glyphosate and Mp*murA* loss-of-function mutants are more sensitive to glyphosate. Furthermore, *A. thaliana* plants overexpressing MpMurA are more resistant to glyphosate. I therefore conclude that that MpMurA does confer glyphosate tolerance, and this is independent of peptidoglycan biosynthesis. I have shown that MpMurA can transfer the enolpyruvyl moiety of PEP to both UDP-GlcNAc and S3P. These data are consistent with the hypothesis that MpMurA contributes to glyphosate tolerance in *M. polymorpha*. However, I also show that MpMurA activity, with the substrates S3P and PEP, is inhibited by glyphosate. This is consistent with glyphosate binding to MpMurA. Therefore, MpMurA may confer glyphosate tolerance through a mechanism analogous to target site overexpression or by binding glyphosate, and in doing so, making it less available to inhibit EPSPS. In either scenario, glyphosate tolerance mediated by MpMurA represents a novel mechanism of herbicide tolerance.

4.6. Supplementary data



Supplementary Fig. S4.1. MurA is very highly conserved among the non-angiosperm streptophytes

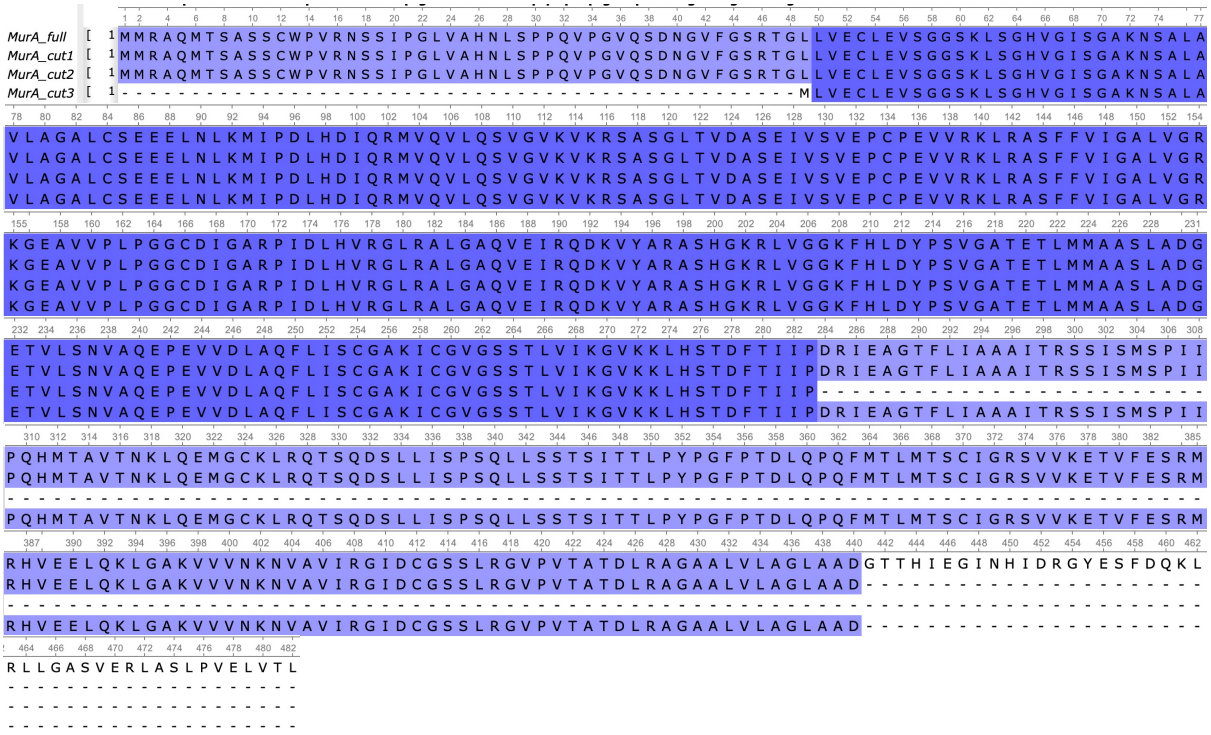
Maximum likelihood tree of MurA amino acid sequences from species from across the Archaeplastida and *E. coli*, rooted to the MurA sequence from *E.coli*. Phylogenetic tree was constructed using PHyML 3.0 (Guindon et al., 2010) using an estimated gamma distribution model and the WAG+G+I model of amino acid substitution. Support values are given from a Chi^2-based aLRT test. Protein sequences were aligned using MAFFT with the L-INS-I algorithm.



Supplementary Fig. S4.2. Overexpression of MpMurA in *A. thaliana* leads to an early bolting phenotype

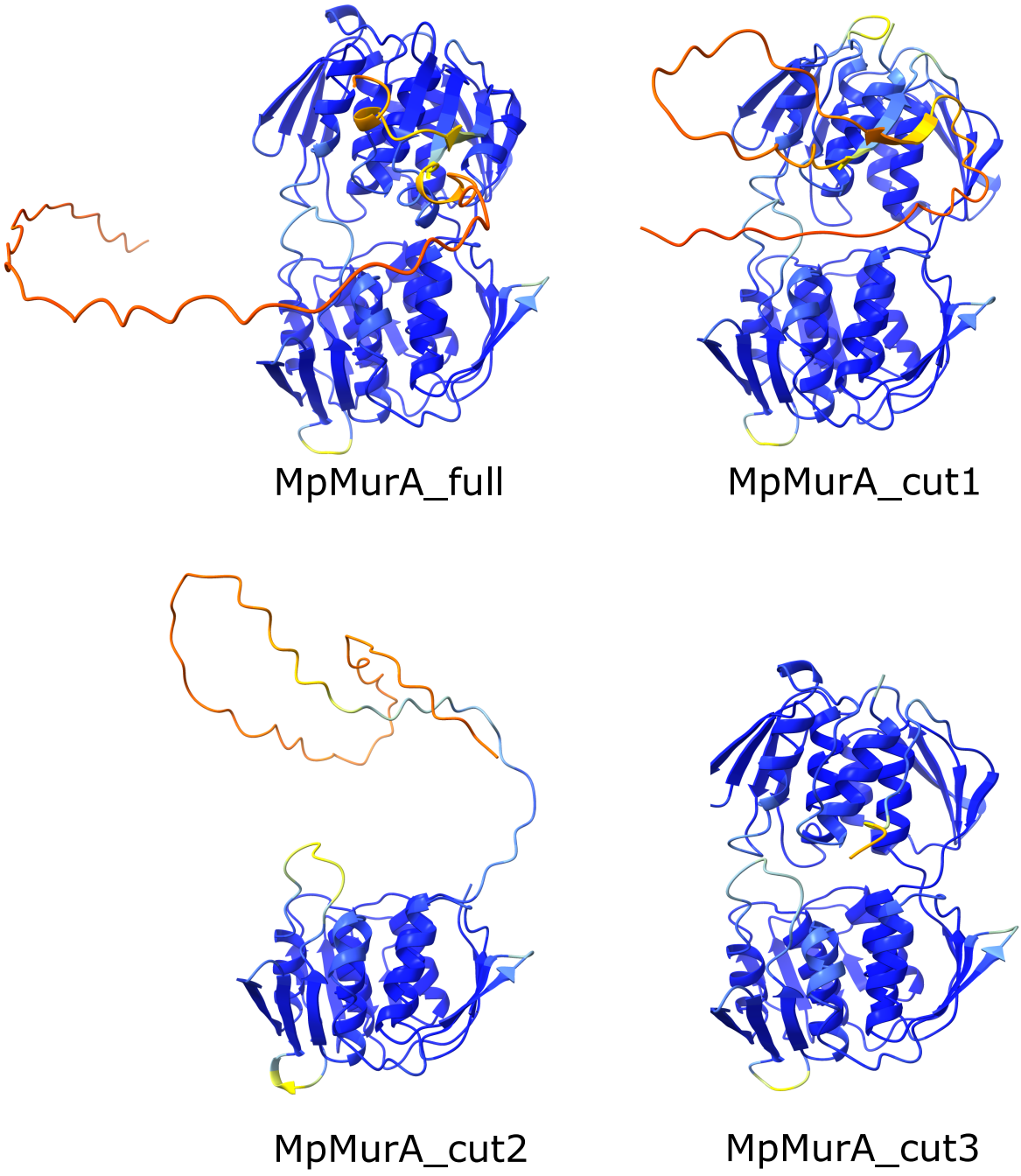
A. Box-and-whisker plots of the days taken for two *A. thaliana* lines expressing MpMurA compared to the wild-type accession Col-0 and a line expressing YFP.

B. Images of wild-type (left) and MpMurA expressing (right) *A. thaliana* plants after 21 days growth on 0 and 1 µM of glyphosate. Images were taken with a Keyence VHX-7000 microscope.



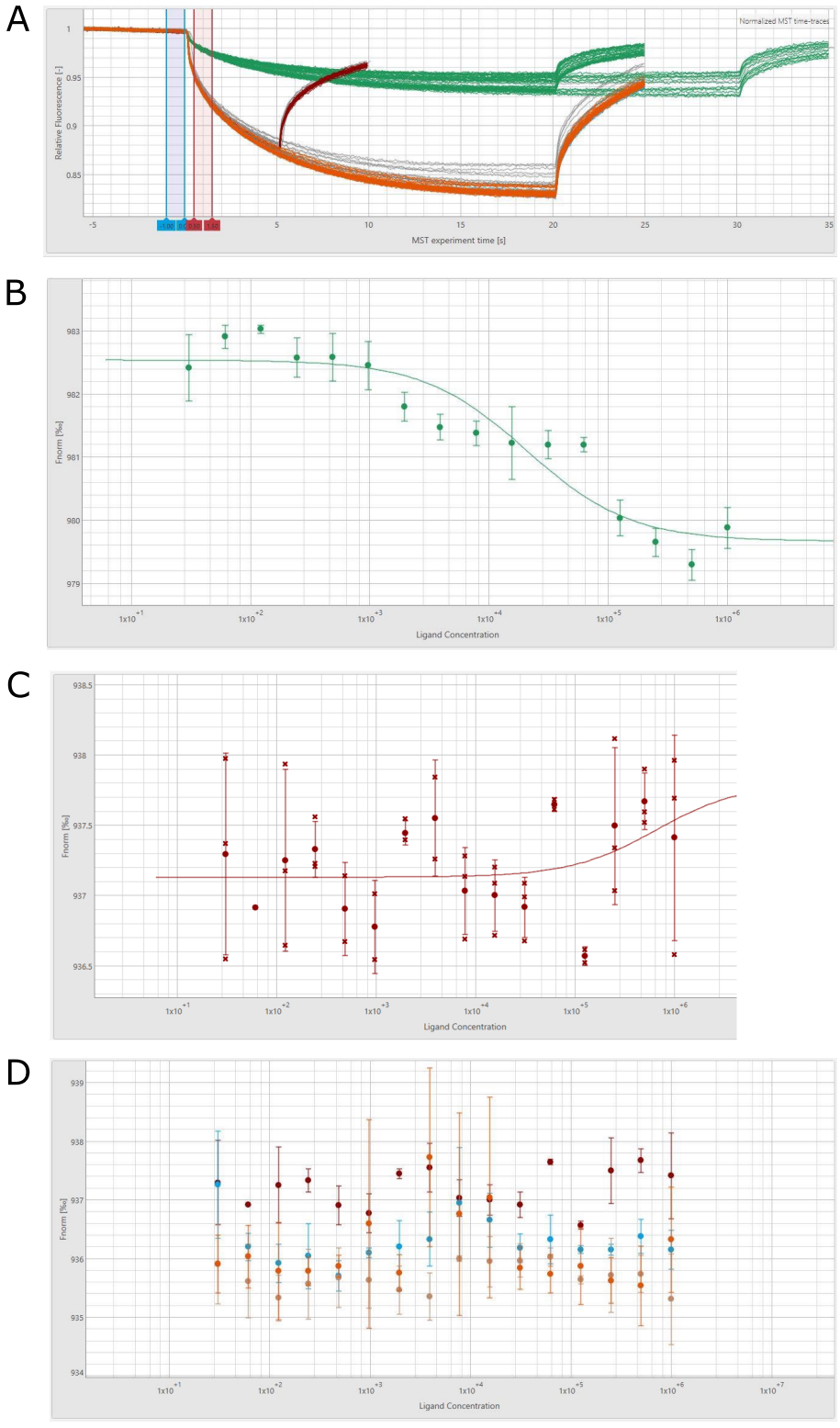
Supplementary Fig. S4.3. Amino acid alignments of the MurA protein and protein fragments expressed in *M. polymorpha*

Alignment was produced using MAFFT via the L-INS-i algorithm.



Supplementary Fig. S4.4. AlphaFold predictions MurA protein and protein fragments expressed in *M. polymorpha*

AlphaFold produced a per residue confidence score (pLDDT) between 0 and 100. Colour denotes confidence in the prediction: Dark blue = Very high (pLDDT > 90); Light blue = Confident (90 > pLDDT > 70); Yellow = Low (70 > pLDDT > 50); Orange = Very low (pLDDT < 50).



Supplementary Fig. S4.5. MST glyphosate binding analysis to MpMurA and MpEPSPS

A. MST traces for all measurements. Depicted in red and blue are the hot and cold regions used for subsequent analysis.

B. Glyphosate dependent normalised fluorescence of MpEPSPS.

C. Glyphosate dependent normalised fluorescence of MpMurA.

D. Normalised fluorescence for 50 nM MurA in presence and absence and with titrations of various small molecules. Red: glyphosate titration, Blue; UDP-GlcNAc titration; Brown: PEP titration in presence of 1mM SP3; Orange: glyphosate titration in presence of 1mM SP3.

Chapter 5: A widespread but distinct metabolomic response occurs following glyphosate-treatment in *Marchantia polymorpha*

5.1. Abstract

Glyphosate is a competitive inhibitor of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which catalyses a key step in the shikimate pathway. Inhibition of EPSPS by glyphosate is proposed to cause plant death by interrupting aromatic amino acid biosynthesis but the accumulation of potentially toxic compounds such as shikimate may also contribute. I demonstrate that shikimate accumulates more in glyphosate-sensitive *A. thaliana* than glyphosate-tolerant *M. polymorpha* following glyphosate treatment. Furthermore, I show that treatment with aromatic amino acids restores growth of *A. thaliana* plants treated with a lethal dose of glyphosate, and treatment with shikimate is slightly toxic. I hypothesise that the metabolic response to glyphosate treatment differs between *M. polymorpha* and *A. thaliana*, as a result of the variation in their glyphosate sensitivity. This difference may be attributed to the low catalytic efficiency of MpEPSPS and MpMurA catalysing the same reaction as MpEPSPS. To investigate this, I performed targeted and non-targeted metabolomic analyses using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Glyphosate treatment did not decrease aromatic amino acid pools but caused an increase in the pools of most amino acids, most notably arginine in *M. polymorpha* and proline in *A. thaliana*. The observed impact of glyphosate treatment on levels of quinic acid, chorismate, and phosphoenolpyruvate in *A. thaliana* and *M. polymorpha* was diverse and complex. My non-targeted metabolomic analysis revealed many metabolic pathways were affected by glyphosate treatment in each species including energy metabolism pathways such as the TCA cycle and glycolysis. The responses of both the TCA cycle and glycolysis differed between *A. thaliana* and *M. polymorpha*, suggesting species-specific metabolic strategies to withstand glyphosate-induced stress. However, an RNA-seq experiment revealed minimal corresponding changes in gene expression in *M. polymorpha*. I conclude that a differential metabolic response to glyphosate treatment in *M. polymorpha*, independent of gene expression changes, may contribute to glyphosate tolerance.

5.2. Introduction

In Chapter 3, I determined that strong glyphosate tolerance in *Marchantia polymorpha* is not due to this plant encoding a tolerant version of the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). In Chapter 4, I have shown that another enzyme in *M. polymorpha*, UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA), also catalyses the reaction which EPSPS catalyses, suggesting that there is redundancy between the functions of the two proteins, which may contribute to glyphosate tolerance. The catalytic efficiency of converting shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) to 5-enolpyruvylshikimate-3-phosphate (EPSP) is lower in MpEPSPS and MpMurA compared to EPSPS enzymes from other wild-type plants. This partly shared function and lower catalytic efficiency likely correlate to a difference in the way *M. polymorpha* regulates the shikimate pathway compared to other plants, which could also contribute to glyphosate tolerance in *M. polymorpha*. This chapter aims to determine if the shikimate pathway is regulated differently in *M. polymorpha* compared to the angiosperm *A. thaliana*. This was done by taking metabolomic measurements following treatment with glyphosate. Furthermore, more widespread metabolomic and transcriptomic changes were searched for in *A. thaliana* and *M. polymorpha* following glyphosate treatment to uncover further potential glyphosate-tolerance mechanisms in *M. polymorpha*.

The shikimate pathway is a critical pathway in plants and is responsible for the biosynthesis of the aromatic amino acids, as well as various secondary metabolites which are involved in plant growth, development, and defence responses (Herrmann and Weaver, 1999; Maeda and Dudareva, 2012; Tzin and Galili, 2010). By inhibiting EPSPS, glyphosate disrupts the biosynthesis of aromatic amino acids, eventually leading to the death of sensitive plants. It is possible for differential regulation of the shikimate pathway to lead to glyphosate resistance. For example, upregulation of EPSPS expression, enhancing its activity, can compensate for glyphosate inhibition – allowing plants to maintain aromatic amino acid production despite glyphosate treatment (Délye et al., 2015; Gaines et al., 2020; Ghanizadeh et al., 2019). Compensation of EPSPS activity in the presence of glyphosate may be carried out by MurA in *M. polymorpha*. There are many examples of where overexpression of EPSPS has evolved in weed populations in the field to confer glyphosate resistance (Alcántara-de la Cruz et al.,

2016; Gaines et al., 2010; Salas et al., 2012; Vila-Aiub et al., 2014). Alternatively, diversion of metabolites away from the shikimate pathway upon glyphosate treatment could also reduce the plant's reliance on EPSPS activity and therefore partly bypass glyphosate inhibition. However, there is little known evidence of such a mechanism. Determining the levels of different components of the shikimate pathway following glyphosate treatment in both the glyphosate sensitive *A. thaliana* and the glyphosate-tolerant *M. polymorpha* may allow me to determine if differential regulation of the shikimate pathway, perhaps associated with a lower catalytic efficiency of MpEPSPS or the activity of MurA, contributes to glyphosate tolerance in *M. polymorpha*.

The low catalytic efficiency of MpEPSPS and the activity of MurA likely result in differential shikimate pathway regulation in *M. polymorpha* compared to species lacking MurA and with more efficient EPSPS enzymes, potentially causing widespread metabolomic differences. Both glyphosate inhibition and a catalytically inefficient EPSPS slow aromatic amino acid synthesis, affecting downstream metabolic pathways. As a compensatory mechanism, *M. polymorpha* may activate alternative metabolic routes, diverting metabolites from the shikimate pathway. Such an adaptation may include utilising or increasing pools of aromatic amino acids through catabolism following glyphosate treatment, a known response in other plants (Maroli et al., 2016). Altered shikimate pathway regulation likely influences intersecting metabolic pathways like the TCA cycle and glycolysis, potentially leading to co-regulation, and secondary metabolite synthesis related to defence responses and growth regulation, which may also contribute to glyphosate tolerance.

M. polymorpha may have evolved other mechanisms to compensate for the low catalytic efficiency of its EPSPS enzyme, which may confer tolerance to glyphosate. Such adaptations may include upregulating alternative metabolic pathways to compensate for the loss of aromatic amino acid synthesis following glyphosate treatment. *M. polymorpha* may reroute carbon and energy flow to bypass the shikimate pathway, sustaining metabolite levels necessary for survival. This rerouting may enhance pathways like the TCA cycle and glycolysis, ensuring energy and intermediates for biosynthetic processes (Mesnage et al., 2016; Zhu et al., 2008). Consequently, *M. polymorpha* likely exhibits distinct metabolic fluxes linked to increased tolerance when treated with glyphosate compared to *A. thaliana*.

Glyphosate treatment is known to induce stress responses in both sensitive and resistant plants, altering secondary metabolism (Eceiza et al., 2022; Zhu et al., 2008). In both cases, secondary metabolites involved in stress responses are upregulated. However, specific metabolic changes may vary between glyphosate-sensitive *A. thaliana* and glyphosate-tolerant *M. polymorpha*, as the latter may have the metabolic tools to deal with inhibition of the shikimate pathway more efficiently.

In this chapter, I have employed metabolomic and transcriptomic approaches to search for differences in the physiological response of *A. thaliana* and *M. polymorpha* to glyphosate treatment, which may give insights into mechanisms of glyphosate tolerance in *M. polymorpha*. I show that the shikimate pathway is inhibited by glyphosate in both *A. thaliana* and *M. polymorpha*, by measuring the accumulation of shikimate. This is consistent with the previous finding that MpEPSPS is glyphosate sensitive. However, other metabolomic responses are different between *A. thaliana* and *M. polymorpha*. This includes pools of intermediates in the shikimate pathway, the TCA cycle and glycolysis. However, I did not find any associated changes in the transcriptome upon glyphosate treatment in *M. polymorpha*, suggesting changes in gene expression are not required to elicit the different metabolic response observed.

5.3. Materials and methods

5.3.1 Plant materials and growth conditions

See Chapter 3 Section 3.4. Plants grown on cellulose discs were transferred using sterile forceps onto plates supplemented with 5 µM glyphosate or control plates after 2 or 7 days.

5.3.2 Growth assays with media supplemented with aromatic amino acids

To test if glyphosate sensitivity in *A. thaliana* can be rescued by supplementation with aromatic amino acids, *A. thaliana* plants were grown on ½ Gamborg medium supplemented with 400 µM of phenylalanine, tryptophan, or tyrosine. *A. thaliana* plants were grown for 3 weeks before imaging with a Berthold Nightowl II LB 983 *In Vivo* Imaging System (Berthold, Bad Wildbad, Germany).

5.3.3 Sensitivity of *M. polymorpha* and *A. thaliana* to shikimate

M. polymorpha gemmae and stratified *A. thaliana* seeds were grown for 14 days on solid ½ Gamborg media (Section 3.3.14) supplemented with varying doses of shikimate before imaging with a Berthold Nightowl II LB 983 *In Vivo* Imaging System (Berthold, Bad Wildbad, Germany).

5.3.4 Measuring shikimate accumulation in plant tissue

4 mm discs were excised from young *A. thaliana* leaves and *M. polymorpha* thallus and placed in separate wells of a 96-well plate pre-filled with 100 µl of 10 mM ammonium phosphate monobasic (adjusted to pH 4.4 with 0.1 N HCl) and 0.1% (v/v) Tween 20 and water. Glyphosate to a final concentration of 0 – 1200 µM was also added before the discs. Plates were sealed and wrapped in Parafilm M and incubated at 23 ℃ in 10-30 µmol m^{-2} s^{-1} white light for 24h. Following incubation, plates were placed at -20 ℃ until the solution was frozen. Samples were thawed at 60 ℃ for 30 min and then 25 µl of 1.24 N HCl was added to each well (for a final concentration of 0.25 N HCl). Plates were incubated at 60 ℃ for a further 15 minutes. 25 µl aliquots were taken from each well and added to fresh wells. Shikimate in these samples was oxidised by adding a solution containing 0.25% (w/v) periodic acid and 0.25% (w/v) sodium m-periodate then incubating at 25 ℃ for 90 mins. Then, 100 µl of 0.6 N sodium hydroxide in 0.22 M sodium sulfite was added to neutralise the

extract and quench the oxidative reaction. Shikimate concentration of each well was determined spectrophotometrically at 380 nm within 30 min using a plate reader (BioTek Synergy H1). Background absorption values measured in wells not exposed to glyphosate were subtracted and final shikimate concentrations calculated from a shikimate standard curve.

5.3.5 Sample extraction for metabolomic analysis

M. polymorpha and *A. thaliana* plants were grown a total of 14 days and exposed to glyphosate treated media for either 2 or 7 days. 150 mg of plant tissue was placed in 2 ml Eppendorf tubes containing a sterile metal bead. Six biological replicates were collected for each treatment group. Samples were frozen in liquid nitrogen and the tissue ground into a powder for 2 minutes using a Tissuelyser (Retsch MM400). Subsequent steps were performed as close to 4 ℃ as possible. To each sample, 500 µl of methanol:acetonitrile:H_2O (2:2:1, v/v) (cooled to -20 ℃) was added, and the mixture homogenized for 2 minutes in the tissuelyser. Samples were incubated at -20 ℃ for 1 hour and centrifuged for 2 minutes at 14,000 rpm. Supernatants were collected and transferred to new Eppendorf tubes and stored at -20 ℃. To the pellet, 400 µl of 80% methanol:water (v/v, stored at -20 ℃) was added, followed by vortexing for 1 minute. Samples were incubated at -20 ℃ for 1 hour and centrifuged for 2 minutes at 14,000 rpm. Supernatants from the first extraction were combined with those from the second extraction for each sample, and samples incubated for 2 hours at -20 ℃. Finally, samples were centrifuged for 10 minutes at 14,000 rpm. 800 µl of supernatant was transferred to new Eppendorf tubes, flash frozen in liquid nitrogen, and stored at -80 ℃ until analysis.

5.3.6 Targeted metabolomic analysis using liquid chromatograph-mass spectrometry

Metabolite extracts from *M. polymorpha* and *A. thaliana* plants were analysed by hydrophilic interaction liquid chromatography (HILIC), directly coupled to liquid chromatography-tandem mass spectrometry (LC-MS/MS). This was performed by the Metabolomics Facility, (Vienna BioCenter Core Facilities). 1 µl of each extract was injected onto a polymeric iHILIC-(P) Classic HPLC column (HILICON, 200 x 2.1 mm; 5 µm) and

respective guard column, and operated at a flow rate of 100 µl/min. A linear gradient (A: acetonitrile; B: 10 mM aqueous ammonium bicarbonate, supplemented with 0.1 µg/ml medronic acid) starting with 25% B and ramping to 90% B in 14 minutes was used for separation. A TSQ Quantiva mass spectrometer (Thermo Fisher Scientific) was used, and SRM transitions used for quantitation can be found in Table 5.1.

Amino acids were analysed in a separate experiment using reversed phase chromatography. 100 µl of each extract was dried in a vacuum centrifuge and dissolved in 100 µl of 0.1% formic acid in water. For each sample, 1 µl was in injected onto a Kinetex (Phenomenex) C18 column (100 Å, 150 x 2.1 mm) with the respective guard column, and employing a 7-minute-long linear gradient from 99% A (1 % acetonitrile, 0.1 % formic acid in water) to 60% B (0.1 % formic acid in acetonitrile) at a flow rate of 80 µl/min. Detection and quantification was done by LC-MS/MS, employing the SRM mode of a TSQ Altis mass spectrometer (Thermo Fisher Scientific), and SRM transitions used for quantitation can be found in Table 5.1.

Authentic standards were used for determining optimal collision energies of the SRM transitions and for validating experimental retention times via standard addition to a pooled quality control sample.

Data interpretation was performed using TraceFinder (Thermo Fisher Scientific).

Compound	Ion mode	SRM transition
glyphosate	negative	*m/z* 168 to *m/z* 150
shikimate-3-phosphate	negative	*m/z* 253 to *m/z* 97
phosphoenolpyruvate	negative	*m/z* 167 to *m/z* 79
5-enolpyruvylshikimate-3-phosphate	negative	*m/z* 323 to *m/z* 79 and *m/z* 323 to *m/z* 97
shikimate	negative	*m/z* 173 to *m/z* 93
quinic acid	negative	*m/z* 191 to *m/z* 83
hexose phosphate	negative	*m/z* 259 to *m/z* 97
pyruvate	negative	*m/z* 87 to *m/z* 43
succinate	negative	*m/z* 117 to *m/z* 73
α-ketoglutarate	negative	*m/z* 145 to *m/z* 101
cis-aconitate	negative	*m/z* 173 to *m/z* 85
dihydroxyacetone phosphate	negative	*m/z* 169 to *m/z* 97
malate	negative	*m/z* 133 to *m/z* 115
citrate	negative	*m/z* 191 to *m/z* 111
diphosphoglycerate	negative	*m/z* 265 to *m/z* 167
glyphosate	positive	*m/z* 170 to *m/z* 88
glycine	positive	*m/z* 76 to *m/z* 30
alanine	positive	*m/z* 90 to *m/z* 44
serine	positive	*m/z* 106 to *m/z* 60
proline	positive	*m/z* 116 to *m/z* 70
valine	positive	*m/z* 118 to *m/z* 72
threonine	positive	*m/z* 120 to *m/z* 74
leucine and isoleucine	positive	*m/z* 132 to *m/z* 86
ornithine	positive	*m/z* 133 to *m/z* 70
asparagine	positive	*m/z* 133 to *m/z* 74
aspartic acid	positive	*m/z* 134 to *m/z* 74
lysine	positive	*m/z* 147 to *m/z* 84
glutamine	positive	*m/z* 147 to *m/z* 130
glutamic acid	positive	*m/z* 148 to *m/z* 84
methionine	positive	*m/z* 150 to *m/z* 133
histidine	positive	*m/z* 156 to *m/z* 110
phenylalanine	positive	*m/z* 166 to *m/z* 133
arginine	positive	*m/z* 175 to *m/z* 70
citrulline	positive	*m/z* 176 to *m/z* 159
tyrosine	positive	*m/z* 182 to *m/z* 136
tryptophan	positive	*m/z* 205 to *m/z* 188

Table 5.1. SRM transitions used for LC-MS/MS quantification.

5.3.7 Non-targeted metabolomic analysis using liquid chromatography-mass spectrometry

To identify untargeted changes in the metabolome of glyphosate-treated *M. polymorpha* and *A. thaliana* plants, LC-MS/MS analysis was conducted using HILIC and reversed phase chromatography (RP) separation methods. Each sample (1 µl) was injected into two different phase systems: a SeQuant ZIC-pHILIC HPLC column (Merck, 100 x 2.1 mm; 5 µm) and a C18-column (Waters, ACQUITY UPLC HSS T3 150 x 2.1; 1.8 µm). For normalization, 6 µl of each sample was combined into a pooled QC sample. Separation was carried out using an Ultimate 3000 HPLC system (Thermo Fisher Scientific, Germany) with a flow rate of 100 µl/min. In the HILIC analysis, a 25-minute gradient from 10% A (acetonitrile - ACN) to 80% B (25 mM ammonium bicarbonate in water) was employed. For reversed phase chromatography (RP), a gradient from 1% to 90% B (A: 0.1% formic acid - FA in water; B: 0.1% FA in ACN) was used. The HPLC system was coupled to the mass spectrometer via electrospray ionization. Metabolites were ionized in polarity switching mode using electrospray ionization, and high-resolution tandem mass spectrometry data were acquired on a Q-Exactive Focus instrument (Thermo Fisher Scientific, Germany) in data-dependent acquisition mode. Compound Discoverer software (Thermo Fisher Scientific) was employed for data processing, utilising in-house and publicly available spectral libraries with a mass accuracy of 3 ppm for precursor masses and 10 ppm for fragment ion masses. Ion counts of each of the identified metabolites were used for subsequent analysis. Log2 fold changes and -log(q value) for each metabolite were calculated. Metabolites identified with a fold change greater than 2 and a -log(q value) less than 2 in *A. thaliana* and *M. polymorpha* after both 2 and 7 days of glyphosate treatment were subject to a pathway analysis using MetaboAnalyst 5.0 (Xia et al., 2009).

5.3.8 RNA-Seq experimental design and RNA extraction and sequencing

Total RNA was extracted from Tak-1 plants at 9 timepoints (0h, 6h, 12h, 18h, 24h, 30h, 36h, 42h, and 48h) after transferring 12-day old plants from cellulose discs to plates treated with 5 µM glyphosate or untreated control plates. This dose is lethal to *A. thaliana* but minimally affects tolerant *M. polymorpha*. Each timepoint had two technical replicate plates and three biological replicates (individual plants), resulting in a total of 54 RNA samples.

Plants collected at each timepoint after transfer onto glyphosate were frozen in liquid nitrogen and stored at -80 ℃. Total RNA was extracted using RNeasy plant mini kits (Qiagen, Hilden, Germany) after grinding with a tissuelyser (Retsch MM400). To remove DNA, DNase treatment was performed using the DNA-free kit by Ambion (Life Technologies) as per the manufacturer's instructions. Subsequently, samples were submitted to the Vienna Biocenter Next Generation Sequencing facility. Samples of sufficient quality proceeded to cDNA library preparation using the NEB poly-A cDNA synthesis kit. Illumina sequencing involved 75 bp paired-end reads and was conducted on the 54 samples with 20-30 million reads per sample in a single lane. Subsequent demultiplexing was carried out on the sequencing results.

5.3.9 Read pre-processing and mapping to the reference transcriptome

Illumina adapters, reads shorter than 25 bp and low quality (with a read quality score lower than 15) were removed using Trimmomatic 0.38 (Bolger et al., 2014). Paired-end reads were interleaved, and ribosomal RNA was removed using SortMeRNA 2.1 (Kopylova et al., 2012) and error corrected using BayesHammer (SPAdes 3.10.0) (Nikolenko et al., 2013) with setting-only-error-correction. Clean reads were aligned (quasi-mapped) to the reference transcriptome (MpTak1v5.1, obtained from https://marchantia.info/download/MpTak1_v5.1r2/) using Salmon (Patro et al., 2017).

5.3.10 Differential expression analysis

Differences in gene expression per transcript between sample treatments were computed using DESeq2 v.1.21.3 (Love et al., 2016) in R. P-values were adjusted using the Benjamini-Hochberg procedure to control for false discovery rate (Benjamini and Hochberg, 1995). Criteria for differential gene expression was an adjusted p-value < 0.05 and a > 2-fold change in expression between sample groups. Expression levels were compared between treated and untreated samples at each timepoint and between treated plants at each timepoint and plants at timepoint 0.

5.4. Results

5.4.1 Glyphosate is taken up in larger quantities in *M. polymorpha* than *A. thaliana* but diminishes at a faster rate

To determine if glyphosate is taken up efficiently into the thallus and leaves of *M. polymorpha* and *A. thaliana* respectively, the quantity of glyphosate was measured using liquid chromatography-tandem mass spectrometry (LC-MS/MS). This was to test if glyphosate tolerance in *M. polymorpha* resulted from insufficient glyphosate reaching photosynthetic tissues. Plants were treated with 5 µM glyphosate for either 2 or 7 days, followed by metabolomic extraction and LC-MS/MS measurements to quantify glyphosate (Fig. 5.1 A and B).

There was a greater quantity of glyphosate measured in *M. polymorpha* compared to *A. thaliana* after both 2 and 7 days of treatment. Glyphosate quantities decreased in both species from day 2 to day 7. Despite *A. thaliana* displaying strong sensitivity to 5 µM glyphosate (Fig. 3.1), it accumulated roughly 36 times less glyphosate in its tissues compared to *M. polymorpha* after 2 days, with values of 133.2 and 4850.4 pg/mg, respectively. After 7 days, *M. polymorpha* still took up approximately 15 times more glyphosate than *A. thaliana*, with values of 253.3043 and 17.1511 pg/mg, respectively. Additionally, glyphosate levels declined more rapidly in *M. polymorpha* compared to *A. thaliana* (Fig. 5.1 C), with about a 19-fold reduction between 2 and 7 days in *M. polymorpha*, versus an 8-fold reduction in *A. thaliana* over the same period. However, none of the glyphosate metabolite aminomethylphosphonic acid (AMPA) was detected in either species.

This is consistent with a differential physiological response to glyphosate treatment in *M. polymorpha* and *A. thaliana*. To test the physiological response of *A. thaliana* and *M. polymorpha* following glyphosate treatment further, the dynamics of shikimate, an upstream compound of the reaction EPSPS catalyses, and an indicator of glyphosate sensitivity, was investigated.

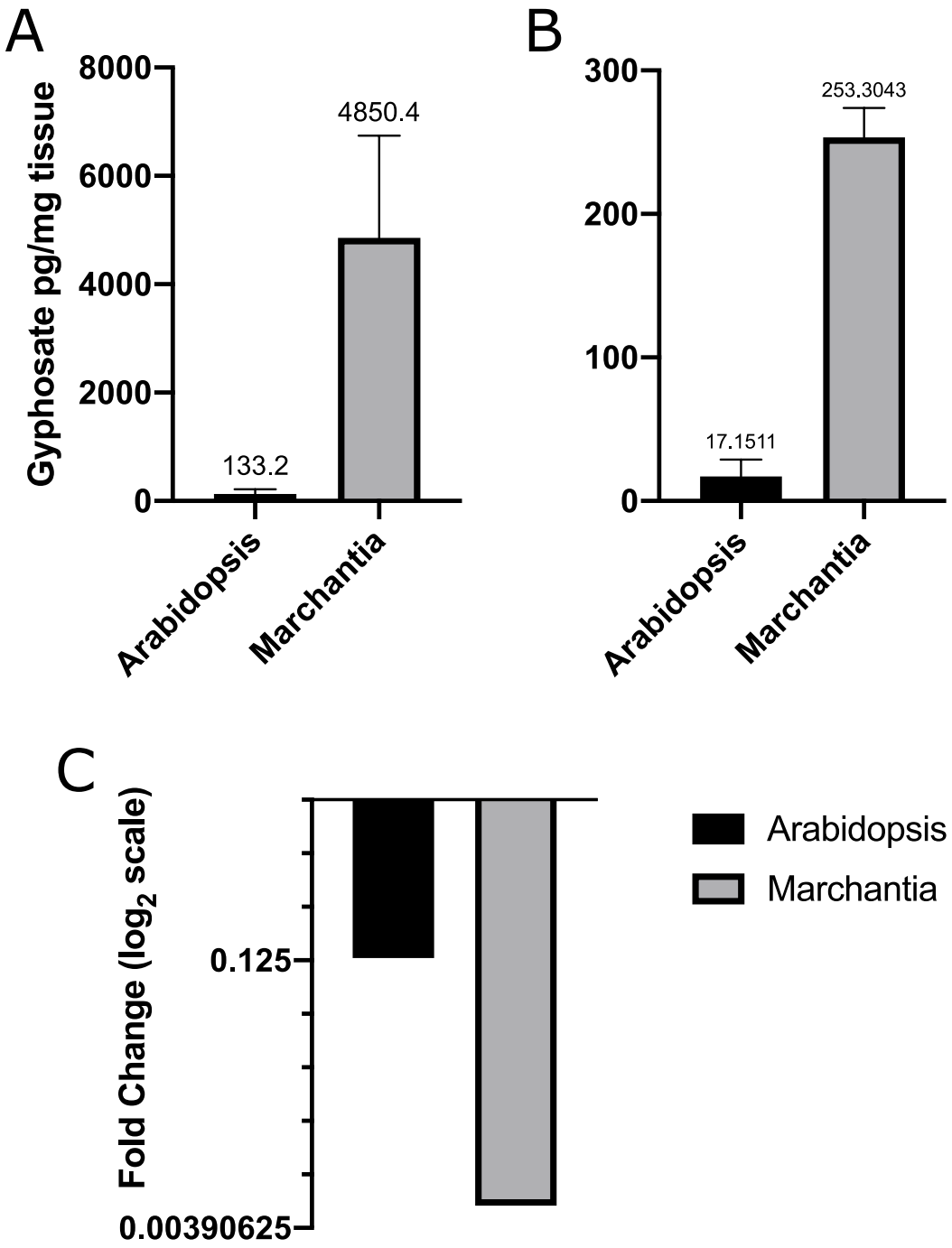


Fig. 5.1. Glyphosate is absorbed in larger quantities into the thallus of *M. polymorpha* than the leaves of *A. thaliana*

Absolute values of glyphosate in plant tissue were measured using LC-MS/MS. Error bars represent SEM.

A. Absolute value measurements of glyphosate in the tissues of 14-day old *A. thaliana* and *M. polymorpha* plants after being exposed to 5 µM of glyphosate for 2 days.

B. Absolute values measurements of glyphosate in the tissues of 14-day old *A. thaliana* and *M. polymorpha* plants after being exposed to 5 µM of glyphosate for 7 days.

C. Fold change of glyphosate found in the tissues of treated *A. thaliana* and *M. polymorpha* plants compared between 2 and 7 days of treatment.

5.4.2 Shikimate accumulates more upon glyphosate treatment in the glyphosate sensitive *A. thaliana* than the glyphosate-tolerant *M. polymorpha*

Glyphosate inhibits the shikimate pathway by binding to EPSPS. It has been widely reported that shikimate accumulates in glyphosate sensitive plants upon treatment with glyphosate (Duke and Powles, 2008; Reddy et al., 2004). Shikimate accumulation in *A. thaliana* and *M. polymorpha* in response to glyphosate treatment was assessed through two methods: spectrophotometric measurement after 24 hours with varying glyphosate doses on living tissue discs (Fig. 5.2 A) and LC-MS/MS analysis after 2 and 7 days of 5 µM glyphosate treatment (Fig. 5.2 B and C).

In living tissues, shikimate accumulates in *A. thaliana* but not *M. polymorpha* within a glyphosate dose range of 0 to 1200 µM. *A. thaliana* initially contains more shikimate (2452.449 µg/g FW) than *M. polymorpha* (1396.043 µg/g FW) at 0 µM glyphosate. In *A. thaliana*, shikimate concentration increases linearly by 3.3-fold to 8093.081 µg/g FW at 1200 µM glyphosate over 24 hours, with a significant deviation from zero (linear regression slope = 4.677, $F = 38.41$, $p < 0.0001$). However, in *M. polymorpha*, there was no increase in shikimate across the same glyphosate concentration range (slope = 0.1059, $F = 0.05755$, $p = 0.8123$). These findings are consistent with *A. thaliana* being glyphosate-sensitive and *M. polymorpha* being glyphosate-tolerant.

To determine more accurately the quantities of shikimate in the tissues of glyphosate-treated *A. thaliana* and *M. polymorpha* LC-MS/MS was employed. The ion counts of shikimate were measured after plants were grown for 2 and 7 days on media supplemented with 5 µM of glyphosate. Inconsistent with measurements taken from living plant discs, LC-MS/MS measurements showed *M. polymorpha* control samples contained more shikimate (9041660 ions; SE = 775561), than *A. thaliana* control samples (260572 ions; SE = 16996). An increase in shikimate was detected in both *A. thaliana* and *M. polymorpha* compared to control (Fig. 5.2 B and C). However, after 2 days, the fold increase in shikimate was 5 times greater in *A. thaliana* (82.3361-fold) compared to *M. polymorpha* (15.7181-fold), a significant difference (Welch's t- test, $p = 0.0042$). After 7 days, both *A. thaliana* and *M. polymorpha* exhibited reduced shikimate fold changes compared to the 2-day treatment,

with values of 8.0749 and 10.4096, respectively. The difference in fold changes between the two species was no longer significant ($p = 0.4778$). Therefore, while *M. polymorpha* is glyphosate tolerant, shikimate accumulates after both 2 and 7 days of glyphosate treatment, suggesting EPSPS inhibition in *M. polymorpha*. However, *A. thaliana* exhibits a higher increase in shikimate after 2 days, consistent with its greater glyphosate sensitivity and EPSPS inhibition. To test the potential impacts of this differential shikimate accumulation on glyphosate MOA, I exogenously applied shikimate to *M. polymorpha* and *A. thaliana*.

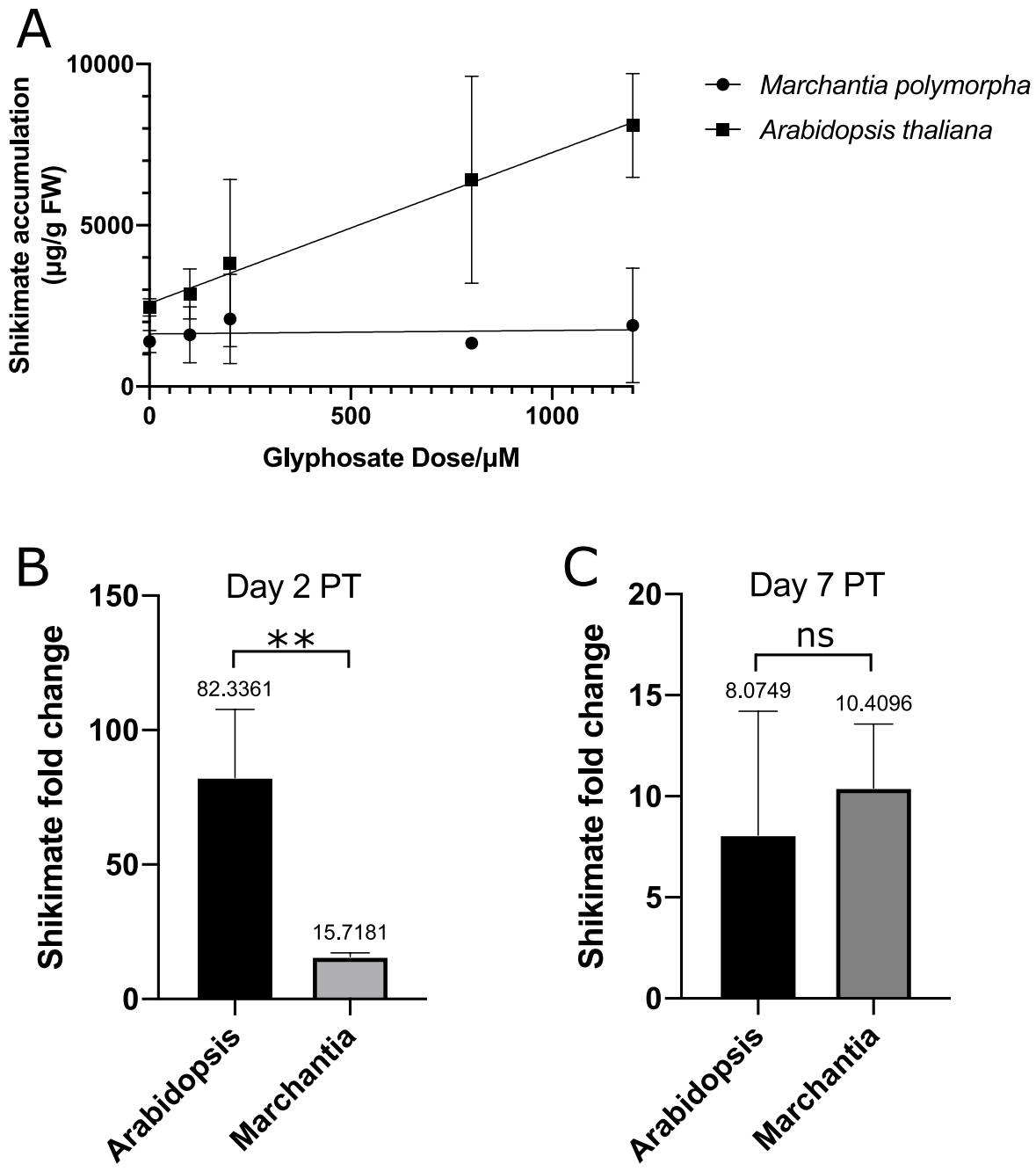


Fig. 5.2. Shikimate accumulates in glyphosate treated plants, but more so in the sensitive *A. thaliana* than tolerant *M. polymorpha*

A. Change in shikimate measured in tissue discs of *A. thaliana* and *M. polymorpha* exposed to increasing doses of glyphosate. Shikimate accumulation was measure spectrophotometrically and µg/g FW calculated from a standard curve. Error bars are SEM and straight lines represent a simple linear regression model fitted to the data.

B. Ion counts of shikimate were measured in *A. thaliana* and *M. polymorpha* plants exposed to glyphosate for 2 days using LC-MS/MS and the fold changes calculated. Error bars represent SEM.

C. Ion counts of shikimate were measured in *A. thaliana* and *M. polymorpha* plants exposed to glyphosate for 7 days using LC-MS/MS and the fold changes calculated. Error bars represent SEM.

Statistical significance in between *A. thaliana* and *M. polymorpha* fold changes in B and C are based on non-parametric Welch's t-test. * = $p < 0.05$, ** = $p < 0.01$, ns = not significant.

5.4.3 Exogenous application of shikimate reduces growth of *A. thaliana* but not *M. polymorpha*

Glyphosate treatment leads to greater shikimate accumulation in the glyphosate-sensitive *A. thaliana* compared to the tolerant *M. polymorpha*. It has been hypothesised that accumulation of shikimate in glyphosate sensitive plants impacts metabolism and may contribute to the toxicity of glyphosate (Nuria de María et al., 2006; Zulet-González et al., 2020). However, there is very little empirical data to support this hypothesis. To test this, shikimate was exogenously applied *M. polymorpha* and *A. thaliana* plants (Fig. 5.3). Plants were grown on media supplemented with varying concentrations of shikimate for 2 weeks and their area measured. There was minimal change in plant area for both *A. thaliana* and *M. polymorpha* across different shikimate doses (Fig. 5.3 A). However, at shikimate doses of 50, 100, and 250 µM, *A. thaliana* exhibited slightly lower plant area relative to control conditions compared to *M. polymorpha*. At 0 µM shikimate, the mean A. thaliana plant area was 96.74 mm^2 (SE = 10.26), but at 250 µM, it significantly reduced to 66.06 mm^2 (SE = 6.505) (student's t-test, p = 0.0224). In contrast, there was no reduction in the mean area of *M. polymorpha* across the same shikimate concentrations (p = 0.4706). The slight toxic effect of shikimate in *A. thaliana* is evident by the reduced area of auto-fluorescing material (Fig. 5.3 C). These findings suggest that in glyphosate-sensitive *A. thaliana*, shikimate has a mild toxic effect, implying a potential role in glyphosate's mode of action. However, this small effect suggests that another factor is likely the primary cause of death in glyphosate-treated plants. Next, the impact of aromatic amino acid starvation on treated plants was assessed.

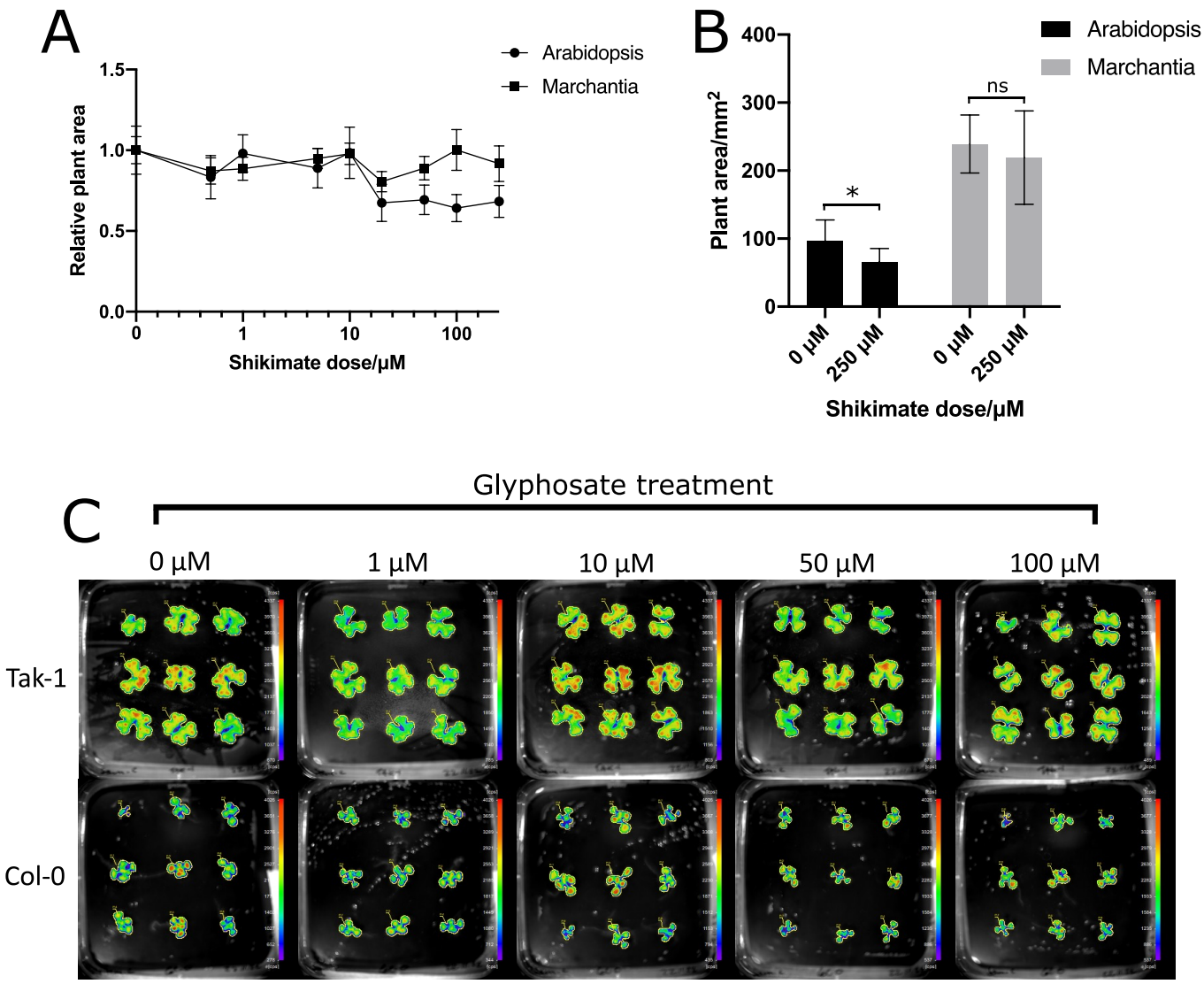


Fig. 5.3. Shikimate slightly inhibits growth in the glyphosate sensitive *A. thaliana* but not in the glyphosate-tolerant *M. polymorpha*

A. The relative plant area of *A. thaliana* and *M. polymorpha* plants was measured and plotted at increasing doses of shikimate. Error bars represent SEM.

B. The mean plant area of *A. thaliana* and *M. polymorpha* plants treated with 0 and 250 µM shikimate was compared. Statistical significance is based on an unpaired t-test (*A. thaliana*, p = 0.0224; *M. polymorpha*, p = 0.4706)

C. Plant area was measured from the total fluorescing area which is taken to be living tissue using the Berthold NightOwl II LB983.

5.4.4 Glyphosate lethality in *A. thaliana* can be rescued with the co-treatment of aromatic amino acids

Glyphosate inhibits the shikimate pathway, which is vital for the biosynthesis of aromatic amino acids (phenylalanine, tryptophan, and tyrosine) in plants. It has been hypothesised that plant death following glyphosate treatment is primarily caused by starvation of these amino acids (Duke and Powles, 2008). To investigate, glyphosate-sensitive *A. thaliana* was treated with a lethal glyphosate dose along with all three aromatic amino acids to assess whether amino acid supplementation could rescue plant lethality (Fig. 5.4). A dose of 10 µM of glyphosate was used as this is lethal to *A. thaliana* plants (Fig. 3.1). The mean area of plants grown on control media without glyphosate or aromatic amino acids was 371.8 mm^2 (SE = 35.87). When *A. thaliana* plants were grown on media supplemented with only 400 µM of each aromatic amino acid, plant area was reduced compared to the control to 52.73 mm^2 (SE = 7.969) (Brown-Forsythe and Welch ANOVA, $p < 0.0001$), indicating that treatment with aromatic amino acids is toxic to *A. thaliana*. Only 2 *A. thaliana* plants of the 27 treated with just 10 µM glyphosate survived, with a mean area of 1.865 mm^2, confirming the lethality of 10 µM of glyphosate. However, when the *A. thaliana* plants were grown on media containing 10 µM glyphosate supplemented with 400 µM of each aromatic amino acid, plant area was restored to the same level as those plants treated with only aromatic amino acids (mean plant area = 80 .85 mm^2; SE = 10.25) ($p = 0.903$). This result is consistent with those plants treated with 10 µM glyphosate alone dying due to a lack of aromatic amino acids as when aromatic amino acids are provided in the media and the same dose of glyphosate applied, the plants no longer die. *M. polymorpha* is tolerant to glyphosate and this may be due to a lack of amino acids starvation in *M. polymorpha* following glyphosate treatment. To test this, a targeted metabolomics approach was employed to measure changes in the pools of aromatic, and all other, amino acids in both *A. thaliana* and *M. polymorpha* following glyphosate treatment.

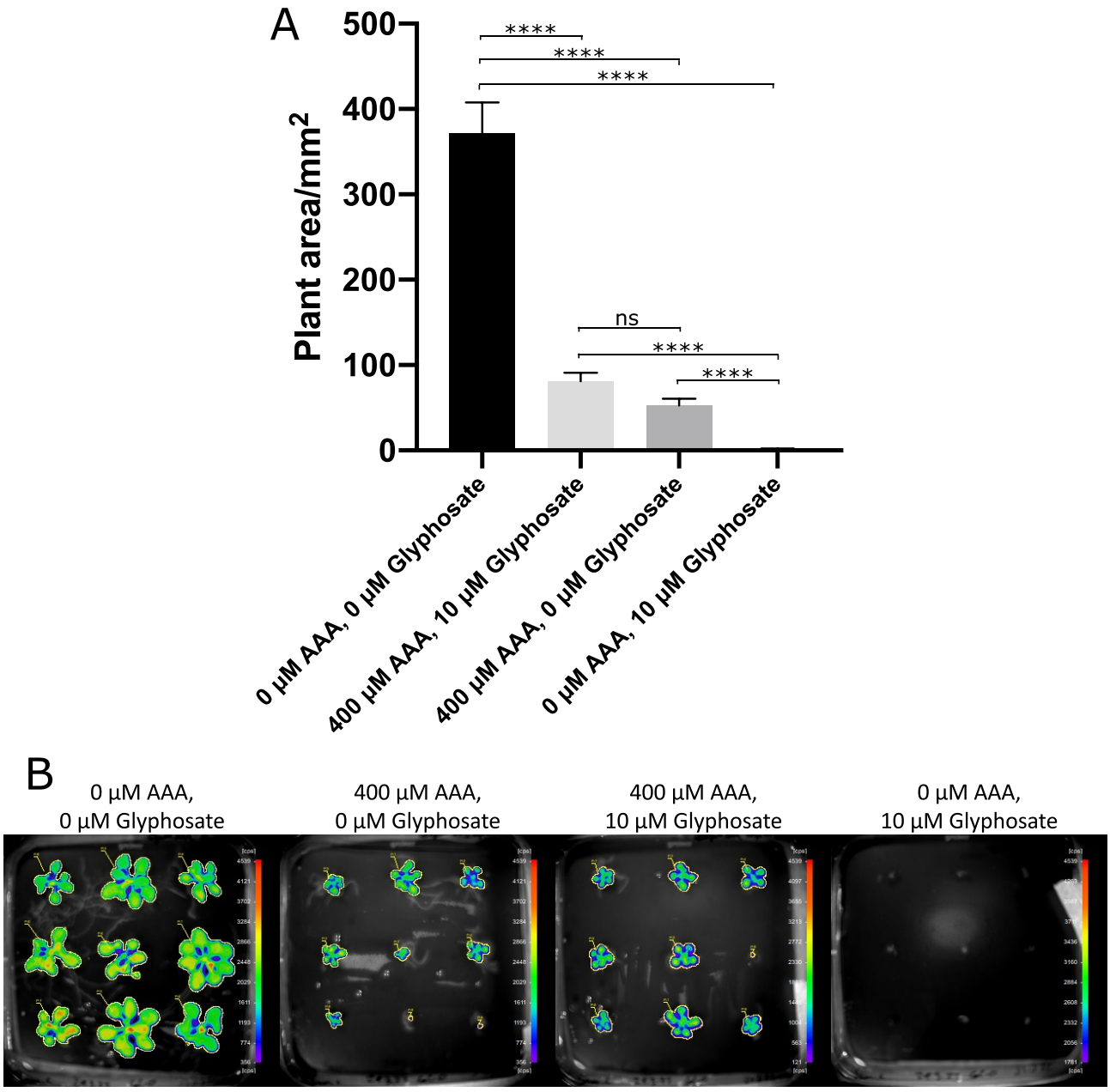


Fig. 5.4. Glyphosate causes death in *A. thaliana* but this is rescued by supplementation with aromatic amino acids

A. Area of Arabidopsis plants treated with 0 µM aromatic amino acids and 0 µM glyphosate, 400 µM aromatic amino acids and 10 µM glyphosate, 400 µM aromatic amino acids and 0 µM glyphosate and 0 µM aromatic amino acids and 10 µM glyphosate. Statistical significance was tested with a Brown-Forsythe and Welch ANOVA test. All treatments were significantly different ($p < 0.0001$) from the others apart from the 400 µM AAA treatments with and without glyphosate compared against each other ($p = 0.903$).

B. Plant area was measured from the total fluorescing area which is taken to be living tissue using the Berthold NightOwl II LB983.

5.4.5 Pools of amino acids are affected differently in *A. thaliana* compared to *M. polymorpha*

Glyphosate likely kills plants by inhibiting aromatic amino acid biosynthesis. To assess the immediate impact of glyphosate treatment on amino acid pools in *A. thaliana* and *M. polymorpha*, I used LC-MS/MS to quantify amino acid levels in plants treated with 5 µM of glyphosate for 2 and 7 days, calculating the fold changes compared to untreated plants (Fig. 5.5).

Glyphosate specifically targets the production of the aromatic amino acids, tyrosine, tryptophan, and phenylalanine. However, there was only a significant difference between the fold change of 2 aromatic amino acids compared between *A. thaliana* and *M. polymorpha* – phenylalanine and tyrosine (t-test adjusted p-value = 0.003789 and 0.002491 respectively) (Fig. 5.5 B). Pools of these amino acids increased in *M. polymorpha*. Phenylalanine and tyrosine were the only amino acids in which there was a different response in *A. thaliana* and *M. polymorpha* after 2 days of glyphosate treatment.

After 7 days of glyphosate treatment, there were more significant differences in amino acid fold changes between *A. thaliana* and *M. polymorpha* (14 differences) compared to 2 days (2 differences). *A. thaliana* exhibited increases in all aromatic amino acids, while *M. polymorpha* showed increased phenylalanine and tyrosine levels but decreased tryptophan levels. Phenylalanine and tryptophan fold changes significantly differed between *A. thaliana* and *M. polymorpha* (adjusted p-values of 0.008798 and 0.008064, respectively). Additional differences included proline and arginine levels; *A. thaliana* exhibited a 12.65-fold increase in proline compared to a 1.42-fold increase in *M. polymorpha*. The most substantial difference was observed in arginine levels, with an 18.42-fold increase in *M. polymorpha* and a 0.56-fold decrease in *A. thaliana* after 7 days of glyphosate treatment, representing a contrasting response to glyphosate treatment regarding this amino acid pool.

In summary, following 2 days of glyphosate treatment, A. thaliana and M. polymorpha displayed distinct aromatic amino acid pool variations, including unexpected rises in phenylalanine and tyrosine in M. polymorpha. After 7 days, both species exhibited more

intricate amino acid pool changes, potentially due to protein breakdown in response to glyphosate-induced stress. The connection between these changes and glyphosate tolerance remains uncertain. To investigate further metabolic factors impacting tolerance, I conducted a metabolomic analysis of shikimate pathway compounds.

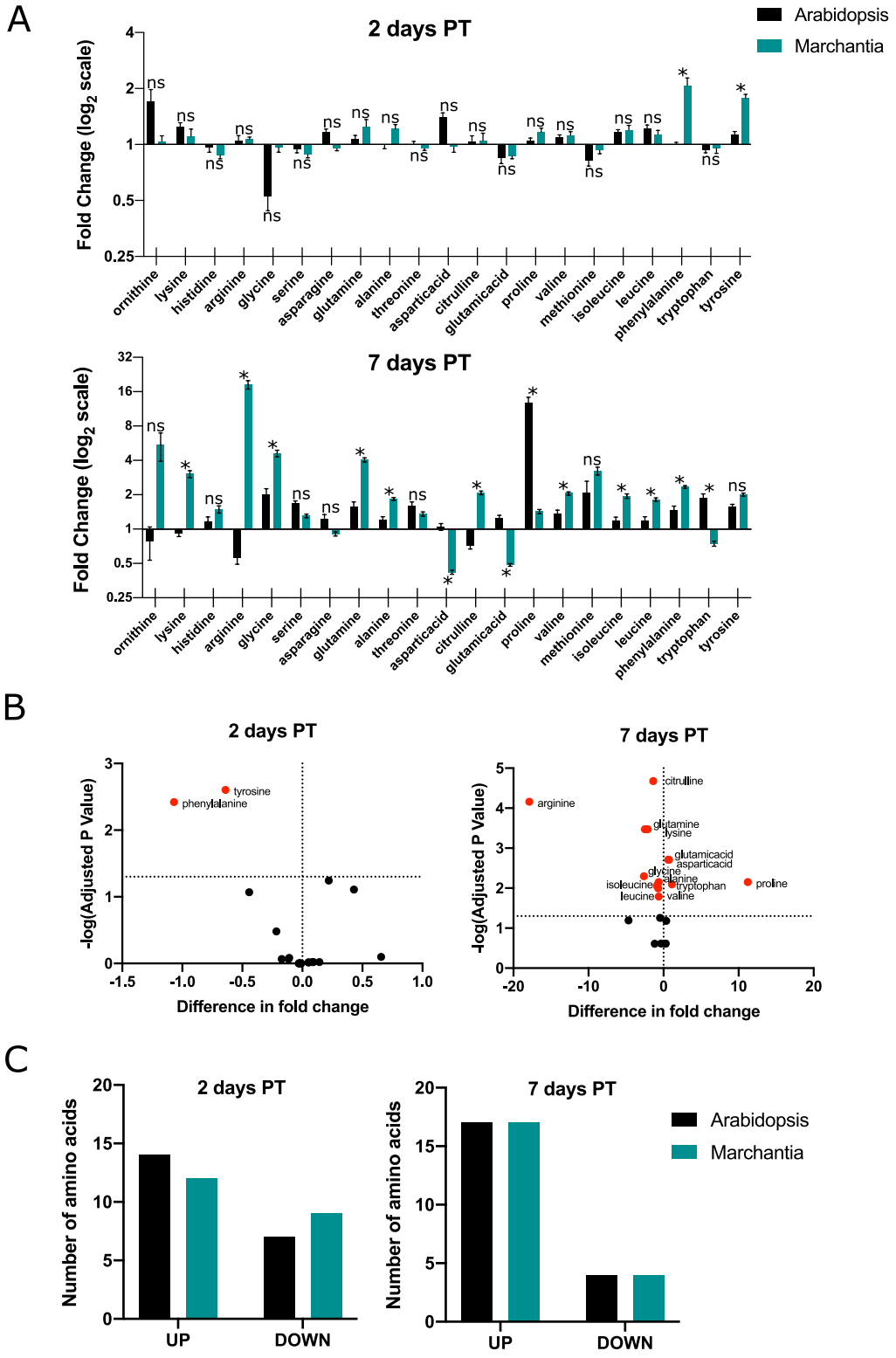


Fig. 5.5. Pools of amino acids are affected differently in the tolerant *M. polymorpha* compared to the sensitive *A. thaliana*

A. Pools of amino acids were measured using LC-MS/MS and fold changes between glyphosate-treated and -untreated plants calculated. Error bars are SEM and significant differences ($p < 0.05$), calculated with a t-test corrected for multiple comparisons using the Holm-Sidak method, between fold change of each amino acid between *A. thaliana* and *M. polymorpha* is shown by *.

B. Volcano plots to summarise the amino acids with significantly different fold changes between *A. thaliana* and *M. polymorpha* after 2 and 7 days of glyphosate treatment.

C. Histograms of the total up and down changes in amino acid pools in *A. thaliana* and *M. polymorpha*.

5.4.6 Quantities of intermediate metabolites of the shikimate pathway are differently affected by glyphosate treatment in *M. polymorpha* and *A. thaliana*

To assess glyphosate's broader impact on the shikimate pathway, I measured quinic acid, phosphoenolpyruvate (PEP), and chorismate levels using targeted LC-MS/MS. I calculated fold changes in ion counts of these compounds between treated and untreated plants after 2 and 7 days of glyphosate treatment and compared these between *A. thaliana* and *M. polymorpha*. Statistical significance was tested for using a Welch's t-test. These compounds were selected for analysis as PEP is an initial shikimate pathway molecule, chorismate is a crucial intermediate in aromatic amino acid biosynthesis, and quinic acid is a potential contributor to glyphosate toxicity (Orcaray et al., 2010; Zulet et al., 2013). Given lower EPSPS inhibition in *M. polymorpha* by glyphosate compared to *A. thaliana*, I anticipate greater accumulation of compounds upstream of EPSPS and less depletion of downstream compounds in *M. polymorpha*. However, predicting PEP and quinic acid quantity changes is challenging due to their involvement in multiple metabolic pathways. Analysing relative changes in these compounds between glyphosate-tolerant *M. polymorpha* and glyphosate-sensitive *A. thaliana* may offer insights into how glyphosate affects the shikimate pathway and potentially other pathways.

After 2 days of glyphosate treatment, significant differences in the fold changes of the three analysed compounds were observed between *A. thaliana* and *M. polymorpha*, indicating distinct metabolic responses. Quinic acid increased in both species, with a greater increase in *A. thaliana* (4.210-fold vs. 2.587-fold, Welch's t-test, $p = 0.0377$). In contrast, changes in chorismate and PEP quantities were opposite. Chorismate increased in *A. thaliana* (5.961-fold) but decreased in *M. polymorpha* (0.521-fold) ($p = 0.0007$), contrary to the expected depletion in *A. thaliana*. PEP increased in *A. thaliana* but decreased in *M. polymorpha* (1.290-fold vs. 0.181-fold, $p = 0.0119$). However, since PEP is involved in multiple metabolic pathways, this change cannot be attributed specifically to the shikimate pathway.

After 7 days of glyphosate treatment, there were minimal changes in quinic acid levels in *A. thaliana* (1.006-fold) but a significant increase in *M. polymorpha* (2.103-fold, $p = 0.0353$). Chorismate levels increased similarly in both species (2.853-fold in *A. thaliana* and 2.842-

fold in *M. polymorpha*) with no significant difference (p = 0.9843). PEP exhibited a consistent pattern with 2 days of glyphosate treatment: an increase in *A. thaliana* and a decrease in *M. polymorpha*. The difference in PEP responses between the two species remained significant (p = 0.0214), with a larger PEP accumulation in *A. thaliana* after 7 days (2.370-fold).

Glyphosate treatment consistently alters the levels of quinic acid, chorismate, and PEP, with differing effects in *A. thaliana* and *M. polymorpha*. This suggests significant and diverse metabolic impacts in both species as a consequence of glyphosate treatment. In order to determine more widespread metabolic responses to glyphosate treatment in these species, a non-targeted metabolomic analysis was conducted.

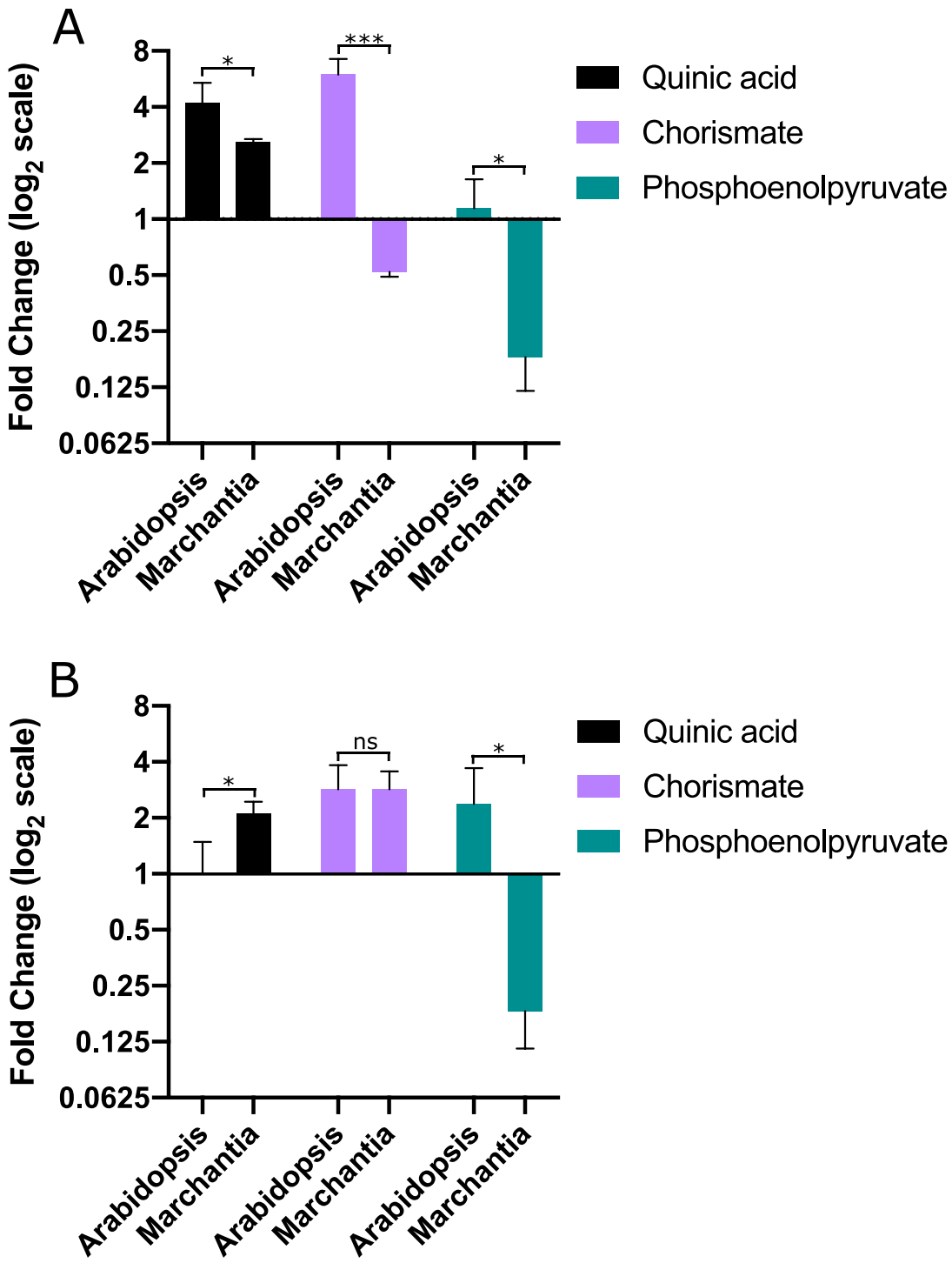


Fig. 5.6. Intermediate metabolites of the shikimate pathway are differentially affected by glyphosate treatment in tolerant and sensitive plants

A. Ion counts of 3 intermediates of the shikimate pathway were measured using LC-MS/MS and the fold changes calculate between glyphosate-treated and -untreated *A. thaliana* and *M. polymorpha* plants after 2 days of glyphosate treatment.

B. Ion counts of 3 intermediates of the shikimate pathway were measured using LC-MS/MS and the fold changes calculated between glyphosate-treated and -untreated *A. thaliana* and *M. polymorpha* plants after 7 days of glyphosate treatment.

Statistical significance between fold changes are based on a Welch's t-test. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, ns = not significant.

5.4.7 Glyphosate treatment causes widespread, but different, metabolic changes in *M. polymorpha* and *A. thaliana*

To investigate global metabolic changes in response to glyphosate treatment in *A. thaliana* and *M. polymorpha*, I conducted non-targeted metabolomic analyses on the same samples that underwent targeted analysis. Two LC-MS/MS separation methods, hydrophilic interaction chromatography (HILIC) and reversed-phase chromatography (RP), were employed to maximize metabolite coverage. Quantitative comparisons were made between glyphosate-treated and untreated samples at both 2 and 7 days post-treatment and volcano plots were generated using log2 fold changes and log(q value) to identify metabolites, the quantity of which significantly altered in response to glyphosate treatment.

Glyphosate treatment induced significant changes in the quantities of metabolites. *M. polymorpha* exhibited more metabolite changes after both 2 and 7 days of treatment compared to *A. thaliana* (Fig. 5.7). In *A. thaliana*, more metabolites increased in quantity than decreased at both time points (Fig. 5.7 A and B). After 2 days, 24 metabolites increased, and 7 decreased, while after 7 days, 53 increased, and 3 decreased. Conversely, *M. polymorpha* showed a more balanced distribution of changes in metabolite quantities after both 2 and 7 days of glyphosate treatment, with 52 metabolites increasing and 51 decreasing after 2 days, and 221 metabolites increasing and 210 decreasing after 7 days (Fig. 5.7 C and D). This suggests that *M. polymorpha* has a broader and more balanced metabolic response to glyphosate treatment compared to *A. thaliana*, despite its tolerance to the herbicide.

To identify metabolites which changed in quantity, common to both *A. thaliana* and *M. polymorpha*, nested Venn diagrams were constructed (Fig. 5.7 E). Quantities of few metabolites consistently changed in both species across various treatment variables and none of the metabolites with altered quantities were common to all treatment variables. After 2 days of glyphosate treatment, there were no metabolites with increased quantity common to both species, increasing to 8 metabolites after 7 days of treatment. A total of 3 metabolites with increased quantity were shared between *A. thaliana* plants treated for 2 days and *M. polymorpha* plants treated for 7 days. For *A. thaliana* plants treated with

glyphosate for 7 days, 1 metabolite which increased in quantity also increased in *M. polymorpha* after both 2 and 7 days of glyphosate treatment. Therefore, few metabolites increased in quantity in response to glyphosate treatment in both *A. thaliana* and *M. polymorpha*, suggesting a differential and dynamic metabolic response in both species.

There were fewer metabolites which decreased in quantity in common between the two species, primarily due there being fewer such metabolites in *A. thaliana* in general. Only 1 metabolite decreased in both species after 7 days of glyphosate treatment. Even within the same species, few metabolites consistently increased or decreased in quantity after both 2 and 7 days of glyphosate treatment. In *A. thaliana*, 9 metabolites increased, while in *M. polymorpha*, 33 increased after both 2 and 7 days. No metabolites decreased after both 2 and 7 days in *A. thaliana*, and only 17 of 210 metabolites decreased in *M. polymorpha* after both time points. These findings indicate that the metabolic response to glyphosate treatment varies between *A. thaliana* and *M. polymorpha* and is dynamic between day 2 and day 7 following treatment in both species.

The non-targeted metabolomics data shows a large and complex response to glyphosate treatment in both *A. thaliana* and *M. polymorpha*. To identify which metabolic pathways the altered metabolites were involved in, a pathway analysis was conducted.

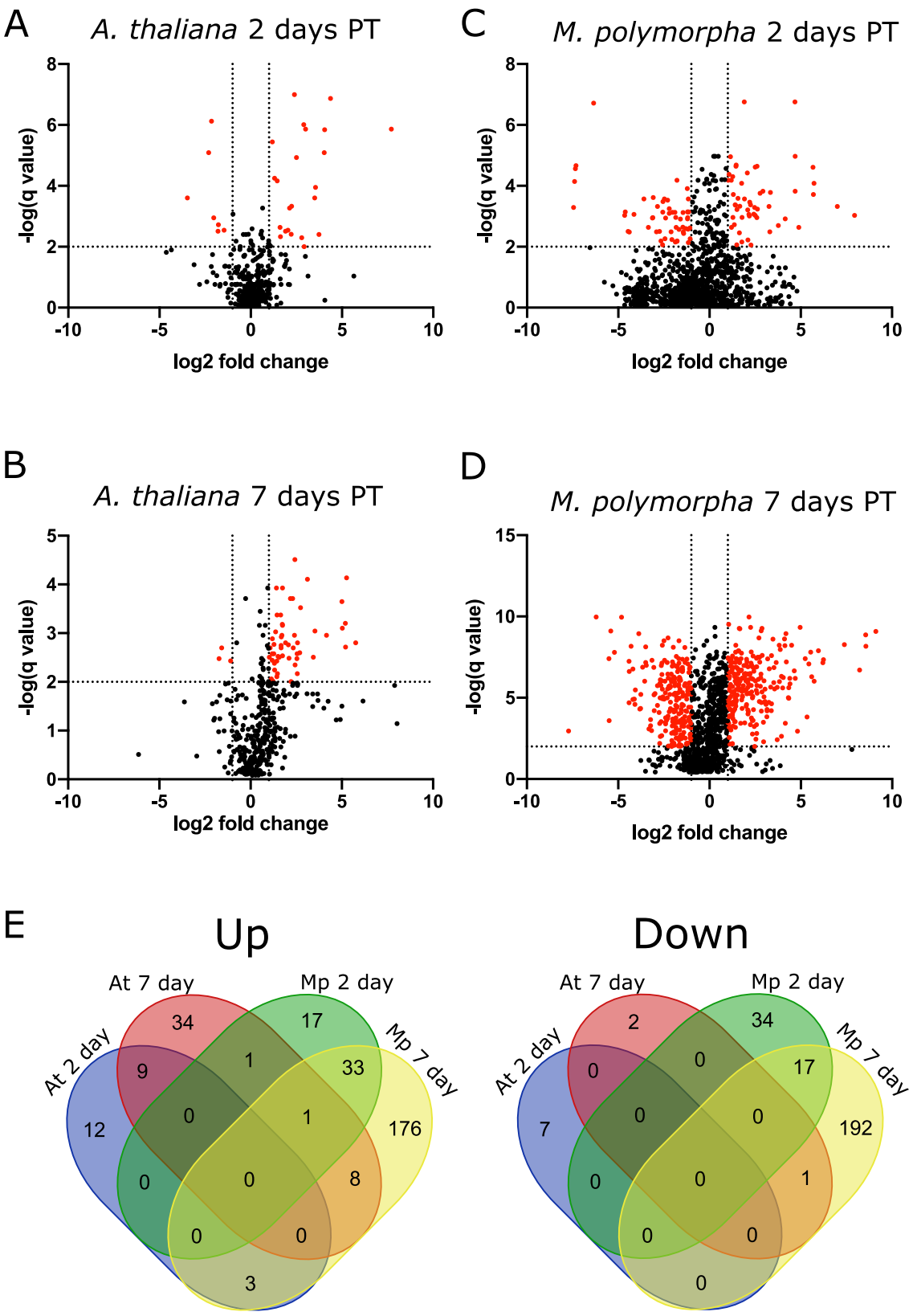


Fig. 5.7. There is a larger metabolomic change upon glyphosate treatment in tolerant *M. polymorpha* than the sensitive *A. thaliana*

A-D. Volcano plot representing significantly changed metabolites in *A. thaliana* (A and B) and *M. polymorpha* (C and D) after 2 (A and C) and 7 (B and D) days of glyphosate treatment from a non-targeted metabolomics analysis using LC-MS/MS. Red dots represent metabolites with a -log(q value) > 2 and minimum 2-fold change compared to untreated samples.

E. Venn diagrams representing the number of metabolites significantly increased or decreased in quantity common to *A. thaliana* and *M. polymorpha* after 2 and 7 days of glyphosate treatment.

5.4.8 Many of the metabolic pathways affected by glyphosate treatment are common between *M. polymorpha* and *A. thaliana*

To determine to which metabolic pathways metabolites with significantly altered quantities, identified in my non-targeted analysis, belonged, a pathway analysis was conducted using MetaboAnalyst 5.0 (Fig. 5.8 A-D). The software identifies links between metabolites based on their associated pathways. Glyphosate treatment had diverse effects on various pathways, which differed between *A. thaliana* and *M. polymorpha* and between 7 and 2 days of glyphosate treatment.

Since glyphosate inhibits a key step in the biosynthesis of the aromatic amino acids, I predicted that this would be one of the pathways most effected by glyphosate treatment. This was the case after 2 days of glyphosate treatment in both *A. thaliana* and *M. polymorpha* (Fig. 5.8 A and C). However, after 7 days of glyphosate treatment, aromatic amino acid biosynthesis was no longer a significantly impacted, suggesting a transient inhibition of the pathway. The analysis also highlighted alterations in other amino acid-related pathways. In *M. polymorpha*, changes occurred in alanine, aspartate, and glutamine metabolism, as well as arginine biosynthesis after both 2 and 7 days of glyphosate treatment. *A. thaliana* displayed impacts on purine, tyrosine, and beta-alanine metabolism, as well as arginine biosynthesis pathways after 2 days (Fig. 5.8 A, C). After 7 days of glyphosate treatment in *A. thaliana*, only the alanine, aspartate, and glutamine metabolism pathways were significantly affected (Fig. 5.8 B). Both species exhibited TCA cycle influence after 7 days. Furthermore, the flavone and flavanol biosynthesis pathway was significantly impacted after 7 days in both species, with more pronounced effect in *A. thaliana*. This impact is expected due to the precursor role of phenylalanine in this pathway.

Among the 31 glyphosate-impacted pathways, 10 were common to both *A. thaliana* and *M. polymorpha* (Fig. 5.8 E). Additionally, 9 and 12 pathways were uniquely affected in *A. thaliana* and *M. polymorpha*, respectively. This indicates that while numerous metabolites showed distinct changes in response to glyphosate treatment in the two plant species (Fig. 5.7), a substantial proportion of the impacted pathways are shared. Some pathways identified, including the TCA cycle, may play a role in glyphosate's mode of action.

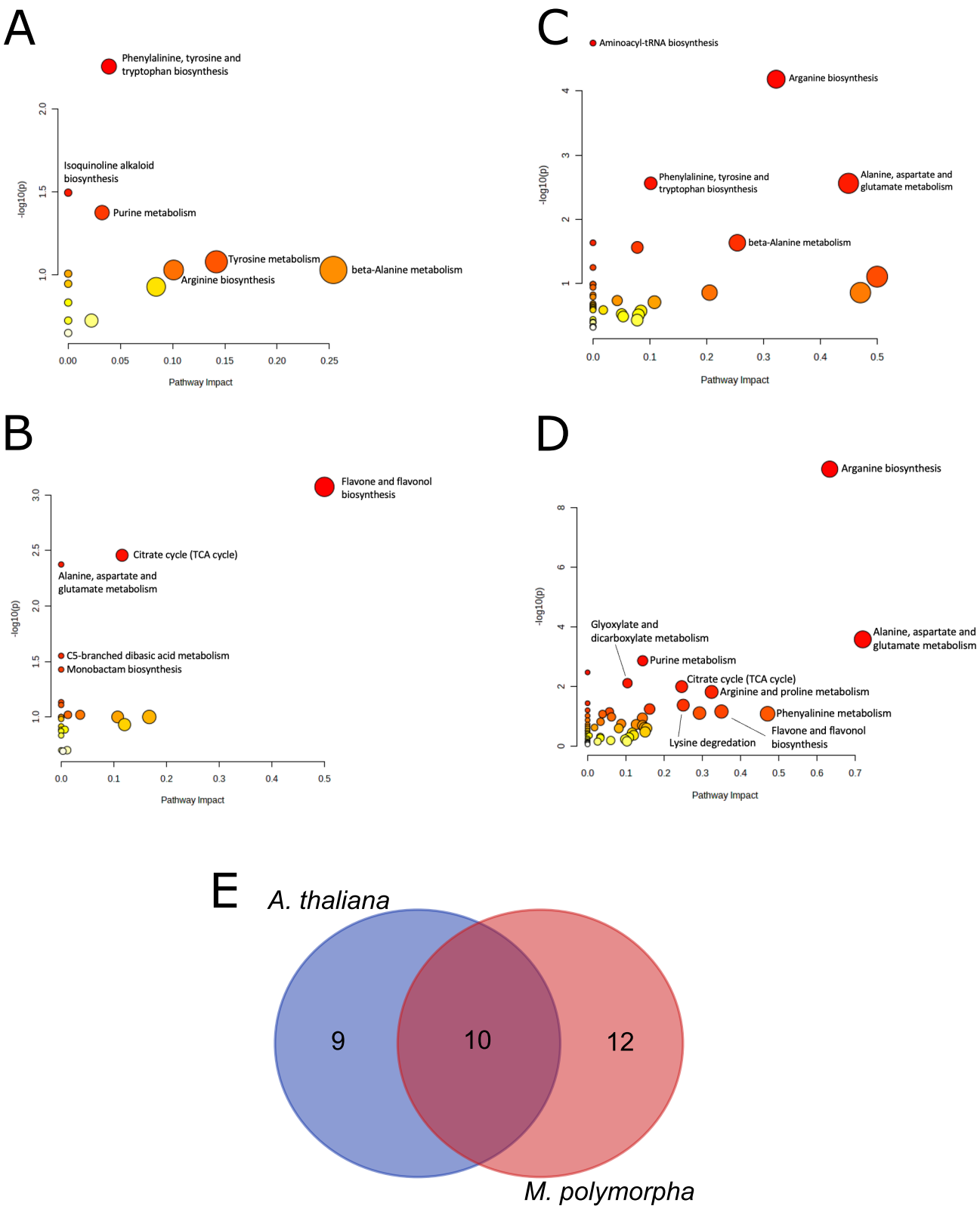


Fig. 5.8. Many metabolic pathways are affected by glyphosate treatment in both tolerant *M. polymorpha* and sensitive *A. thaliana* plants

A-D. Results of a pathway analysis of the significantly increased or decreased metabolites in (A) *A. thaliana* after 2 days glyphosate treatment; (B) *A. thaliana* after 7 days glyphosate treatment; (C) *M. polymorpha* after 2 days glyphosate treatment; (D) *M. polymorpha* after 7 days glyphosate treatment. Size of the points is relative to the number of metabolites belonging to that pathway and darker colours depict higher levels of statistical significance. E. Venn diagram of depicting the number of pathways shared and unique affected across all treatments in *A. thaliana* and *M. polymorpha*.

5.4.9 There are large and opposite impacts on the TCA cycle in *A. thaliana* and *M. polymorpha* after 7 days of glyphosate treatment

The TCA cycle is essential for extracting energy from organic compounds in plants. Through non-targeted metabolomics, I found that glyphosate treatment impacted the TCA cycle more in the sensitive plant *A. thaliana* than *M. polymorpha* after 7 days. Using targeted LC-MS/MS on the same samples as previous analyses, I quantified the levels of 5 TCA cycle intermediates (citrate, α-ketoglutarate, succinate, malate, and cis-Aconitate) to assess the impact of glyphosate treatment on this pathway. Fold changes between untreated and treated samples were calculated for both 2 and 7 days of glyphosate treatment and compared between *A. thaliana* and *M. polymorpha* (Fig. 5.9). A Welch's t-test was employed to test for statistical significance between fold changes. After 2 days, the TCA cycle remained largely stable in *A. thaliana*, with significant changes in 3 compounds in *M. polymorpha*. However, after 7 days of glyphosate treatment, substantial and often opposing changes occurred in these measured compounds in both species.

After 2 days of glyphosate treatment, no significant changes were observed in the quantity of TCA cycle compounds in *A. thaliana* (statistical significance calculated from a 1 sample t and Wilcoxon test). In contrast, in *M. polymorpha*, citrate decreased by 0.8065-fold ($p = 0.0367$), α-ketoglutarate decreased by 0.3269-fold ($p = 0.0001$), and malate decreased by 0.6453-fold ($p = 0.0202$). Thus, after 2 days of glyphosate treatment, the TCA cycle in *A. thaliana* remained unaffected, while in *M. polymorpha*, three of the five tested intermediates decreased significantly. Additionally, cis-aconitate increased substantially (5.818-fold) in *M. polymorpha*, though this increase was not deemed significant due to sample variance ($p = 0.0890$).

After 7 days of glyphosate treatment, significant changes occurred in most measured TCA cycle intermediates in both *A. thaliana* and *M. polymorpha* (Fig. 5.9 B). However, in *A. thaliana*, α-ketoglutarate and cis-aconitate did not significantly differ from 1 ($p = 0.4149$ and $p = 0.0562$ respectively) after the treatment. In *M. polymorpha*, citrate showed no significant change after 7 days ($p = 0.2316$), indicating a rapid and transient alteration in citrate levels. In *A. thaliana*, all 5 TCA cycle intermediates had mean fold changes greater than 1,

suggesting accumulation after 7 days of glyphosate treatment. Conversely, in *M. polymorpha*, only cis-aconitate had a fold change greater than 1 (8.822-fold), with all other measured TCA cycle compounds decreasing. α-ketoglutarate showed no statistically significant difference between *A. thaliana* and *M. polymorpha* but exhibited a significant decrease in *M. polymorpha*. There were opposite trends in the changes of citrate, succinate, and malate between *A. thaliana* and *M. polymorpha* ($p < 0.0001$, $p = 0.0052$, and $p = 0.0008$ respectively). Cis-aconitate levels significantly accumulated after 7 days of glyphosate treatment in both *A. thaliana* (2.931-fold) and *M. polymorpha* (8.822-fold), with the increase in *M. polymorpha* being significantly greater ($p = 0.0004$).

In summary, the TCA cycle, which is not directly inhibited by glyphosate, is affected by glyphosate treatment. This impact differs between *A. thaliana* and *M. polymorpha* and varies between 2- and 7-days of treatment. My analysis reveals opposite effects on the TCA cycle in glyphosate-sensitive *A. thaliana* compared to glyphosate-tolerant *M. polymorpha*, which could be due to differences in the general stress response of each species or specific TCA cycle effects from glyphosate treatment. To assess if these trends extend to other major metabolic energy pathways, components of glycolysis were also measured.

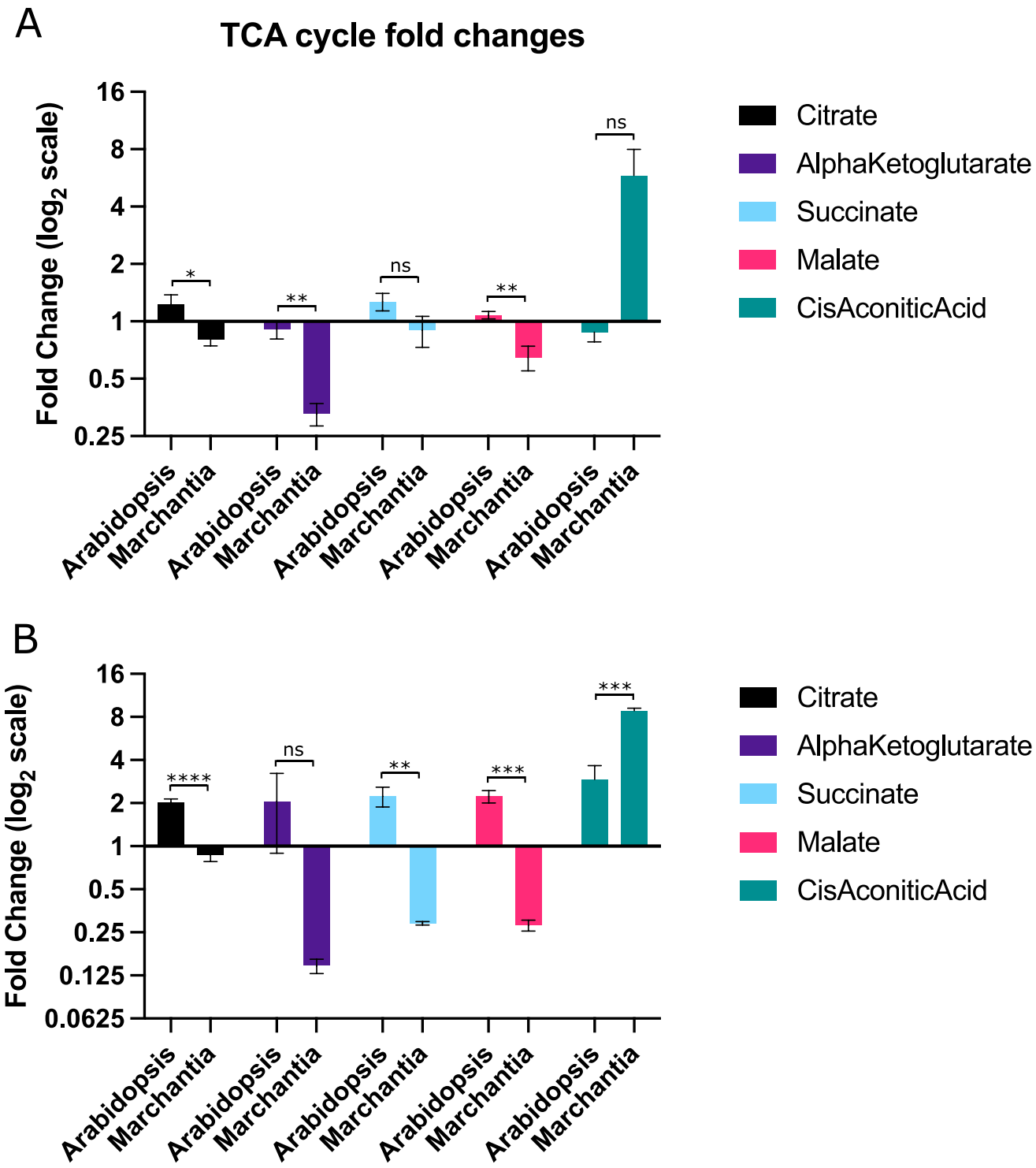


Fig. 5.9. The TCA cycle is differentially affected by glyphosate treatment in the tolerant *M. polymorpha* and sensitive *A. thaliana*

A. Fold changes from glyphosate treated and untreated plants of intermediates of the TCA cycle measured with LC-MS/MS after 2 days of treatment.

B. Fold changes from glyphosate treated and untreated plants of intermediates of the TCA cycle measured with LC-MS/MS after 7 days of treatment.

Statistical significance between the fold changes for *A. thaliana* and *M. polymorpha* for each compound are based on a Welch's t-test. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, ns = not significant.

5.4.10 Many intermediates of glycolysis decrease in *M. polymorpha* upon glyphosate treatment but increase in *A. thaliana*

Glycolysis is the central metabolic pathways that breaks down glucose, providing energy for cellular processes, and is likely impacted by glyphosate treatment. *A. thaliana* is glyphosate sensitivity and *M. polymorpha* is tolerant, suggest differing impacts on glycolysis. I quantified hexose-6-phosphate, DHAP, PEP, diphosphoglyceric acid, and pyruvate via LC-MS/MS in glyphosate-treated and untreated *A. thaliana* and *M. polymorpha* calculated the fold changes after 2 and 7 days (Fig. 5.10). These were compared between *A. thaliana* and *M. polymorpha* and a Welch's t-test was employed to test for statistical significance. More significant changes occurred after 7 days in both species: *M. polymorpha* reduced glycolytic intermediates, while *A. thaliana* showed variable changes.

After 2 days of glyphosate treatment, no significant changes occurred glycolytic intermediates in *A. thaliana* (Fig. 5.10 A) (statistical significance calculated from a 1 sample t and Wilcoxon test). However, in *M. polymorpha*, 2 of the 5 intermediates deviated significantly from 1: PEP decreased by 0.6401-fold ($p = 0.0044$), and diphosphoglyceric acid decreased by 0.4985-fold ($p = 0.0012$). This indicates no significant impact on measured glycolytic intermediates in *A. thaliana*, while *M. polymorpha* shows a significant decrease in 2 intermediates. Hexose-6-phosphate, DHAP, and pyruvate remained unchanged in both species after 2 days of glyphosate treatment.

After 7 days of glyphosate treatment, there were three unchanged glycolytic intermediates in *A. thaliana*: PEP, diphosphoglyceric acid, and pyruvate ($p = 0.1195$, $p = 0.0552$, and $p = 0.4331$ respectively). In *M. polymorpha*, only pyruvate remained unchanged (0.8782-fold, $p = 0.0756$). Conversely, *M. polymorpha* experienced significant decreases in all other measured intermediates: hexose-6-phosphate (0.3443-fold, $p = 0.0002$), DHAP (0.2607-fold, $p < 0.0001$), PEP (0.1396-fold, $p = 0.0001$), and diphosphoglyceric acid (0.1377-fold, $p < 0.0001$). *A. thaliana* also showed significant decreases in hexose-6-phosphate (0.6268-fold, $p = 0.0135$) and DHAP (0.3695-fold, $p = 0.0036$). Hexose-6-phosphate decreased significantly more in *M. polymorpha* than in *A. thaliana* (Welch's t-test, $p = 0.0303$), while DHAP decrease was similar between the two species ($p = 0.3542$). PEP and diphosphoglyceric acid displayed

opposite trends: fold changes were greater than 1 in *A. thaliana* but less than 1 in *M. polymorpha* ($p = 0.0184$ and 0.0138 respectively).

In summary, significant changes in glycolysis intermediates were observed after glyphosate treatment, especially after 7 days and particularly in *M. polymorpha*. Despite glyphosate tolerance at the whole plant level, *M. polymorpha* exhibits substantial glyphosate-induced effects on glycolysis. These effects are generally more pronounced in *M. polymorpha* and sometimes oppose those in glyphosate-sensitive *A. thaliana*. However, the implications of these metabolic differences on glyphosate tolerance remain unclear. To explore potential gene expression changes in *M. polymorpha* that may explain these metabolic differences or respond to metabolic shifts, the transcriptome of glyphosate-treated *M. polymorpha* was sequenced.

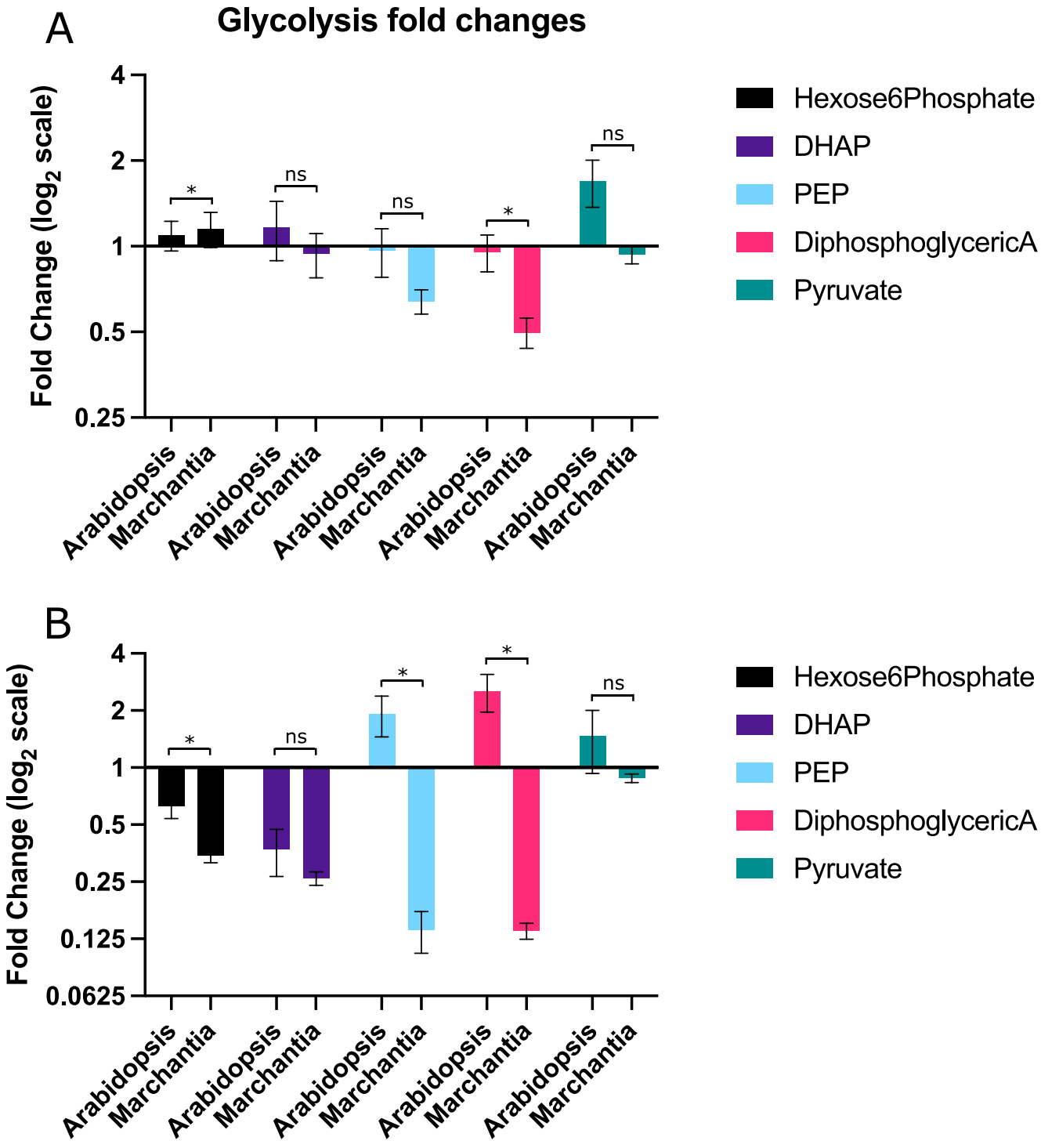


Fig. 5.10. Glycolysis is differentially affected by glyphosate treatment in the tolerant *M. polymorpha* and sensitive *A. thaliana*

A. Fold changes from glyphosate treated and untreated plants of intermediates of glycolysis measured with LC-MS/MS after 2 days of treatment.

B. Fold changes from glyphosate treated and untreated plants of intermediates of glycolysis measured with LC-MS/MS after 7 days of treatment.

Statistical significance between the fold changes for *A. thaliana* and *M. polymorpha* for each compound are based on a Welch's t-test. * = $p < 0.05$, ns = not significant.

5.4.11 There are very few transcriptome changes in *M. polymorpha* following glyphosate treatment

My metabolomic analysis of glyphosate-treated *M. polymorpha* and *A. thaliana* plants revealed widespread changes to the metabolome, even in the glyphosate-tolerant *M. polymorpha*. To test if the impacts on the metabolome of glyphosate treatment on *M. polymorpha* were driving or responding to transcriptome changes over the same time period, an RNA-seq experiment was carried out on Tak-1 *M. polymorpha* plants over a 48-hour period following glyphosate treatment. Plants were transferred onto plates with 0 µM or 5 µM of glyphosate, with samples collected at 0, 6, 12, 18, 24, 30, 36, 42 and 48 hours. cDNA libraries and the 75 bp paired-end reads were sequenced by Illumina sequencing. Gene expressions were compared between glyphosate-treated and -untreated samples for each timepoint and between treated plants and plants at timepoint 0.

A total of 26 genes exhibited differential expression patterns across all timepoints and treatments. Among them, 23 genes showed differential expression in untreated plants at various timepoints compared to untreated samples at timepoint 0 (Supplementary Table S5.1). However, minimal changes were observed in glyphosate-treated *M. polymorpha* plants when compared to untreated ones (Table 5.2). Only three genes displayed differential expression between glyphosate-treated and untreated samples across all timepoints. The expression of Mp2g12020 decreased at 18, 24, and 30 hours after transferring plants to glyphosate-supplemented media, with respective log2 fold changes of 0.6589, 0.6799, and 0.6512 (adjusted p-values = 0.0267, 0.0134, and 0.0323). Furthermore, at 48 hours following the transfer to glyphosate media, two genes, Mp7g02900 and Mp8g00430, exhibited increased expression by 1.4943- and 1.1559-fold, respectively (adjusted p-values = 0.0076 and 0.0110). The functions of these three genes are unknown.

In summary, the transcriptomic analysis of glyphosate-treated *M. polymorpha* indicates minimal changes in gene expression due to both glyphosate treatment and the previously observed alterations to the metabolome. None of the differentially expressed genes have known functions. Consequently, over this time course, transcriptome changes cannot

account for the metabolic distinctions between *M. polymorpha* and *A. thaliana* or elucidate the mechanism underlying glyphosate tolerance in *M. polymorpha*.

A

Timepoint PT	0h	6h	12h	18h	24h	30h	36h	42h	48h
Differentially expressed genes	0	0	0	1	1	1	0	0	2

B

Locus name	Gene function	Timepoint differentially expressed	Log2 Fold Change	Adjusted p-value
Mp2g12020	unknown	18h	0.6589	0.0267
Mp2g12020	unknown	24h	0.6799	0.0134
Mp2g12020	unknown	30h	0.6512	0.0323
Mp7g02900	unknown	48h	1.4943	0.0076
Mp8g00430	unknown	48h	1.1559	0.0110

Table 5.2. Very few genes are differentially expressed upon glyphosate treatment in *M. polymorpha*
A. Number of differentially expressed genes in glyphosate-treated samples compared to the untreated samples at the corresponding timepoint.
B. Overview of the differentially expressed genes linked to glyphosate treatment with the log2 fold change and adjusted p-value given.

5.5. Discussion

In this chapter, I have determined the changes in the metabolome of *A. thaliana* and *M. polymorpha* following glyphosate treatment. I have identified a variety of different responses, based on different metabolite fluxes, in the shikimate pathway, tricarboxylic acid (TCA) cycle and glycolysis. I have shown that shikimate accumulates more in glyphosate-treated *A. thaliana* than *M. polymorpha* and exogenous shikimate treatment is slightly toxic to *A. thaliana* and not *M. polymorpha*. This is consistent with shikimate accumulation in *A. thaliana* contributing to the lethal effect of glyphosate treatment. Furthermore, I have determined that there are differences in the regulation of pools of amino acids following glyphosate treatment in *A. thaliana* and *M. polymorpha*. I showed that supplementation with aromatic amino acids could rescue the lethality of glyphosate treatment in *A. thaliana*, suggesting aromatic amino acid starvation is the cause of plant death. However, starvation of aromatic amino acids was not detected using LC-MS/MS. This may reflect the complex regulation of aromatic amino acid biosynthesis in plants. This complex regulation is supported by my finding of widespread changes to the metabolome of both *A. thaliana* and *M. polymorpha* following glyphosate treatment which changed over time. This suggests glyphosate treatment elicits a cascade of metabolic changes in plants, some of which may contribute to plant death. These metabolic responses were often different between the species and may be associated with the difference in glyphosate tolerance in these two species or the role of MurA catalysing the reaction of EPSPS in *M. polymorpha*. There was no clear transcriptional response to glyphosate treatment in *M. polymorpha*. This suggests that the large-scale metabolic changes observed in *M. polymorpha* are not due to transcriptional regulation, suggesting post-transcriptional or translational regulatory mechanisms may contribute to glyphosate tolerance in *M. polymorpha*. These may represent important, undiscovered mechanisms which are responsible for herbicide resistance in populations of weeds in the field.

Glyphosate tolerance in *M. polymorpha* may be due to limited glyphosate uptake. Using liquid chromatography-tandem mass spectrometry (LC-MS/MS), I showed that glyphosate was taken up into the tissues of *M. polymorpha*. In fact, 36-fold to 15-fold more glyphosate was found in the tissues of *M. polymorpha* than *A. thaliana*. The higher concentration of

glyphosate within *M. polymorpha* compared to *A. thaliana*, suggests a many times higher level of glyphosate tolerance in *M. polymorpha* compared to *A. thaliana* than is observed at the whole plant level.

Using a targeted metabolomic approach I measured the quantity of various intermediates of the shikimate pathway, using LC-MS/MS, to investigate how the shikimate pathway is regulated in both *A. thaliana* and *M. polymorpha* in response to glyphosate treatment. An increase in shikimate was observed in both species which suggests that glyphosate inhibits EPSPS in both plants. Despite the tolerance of *M. polymorpha* to glyphosate, measured accumulation of shikimate demonstrates that EPSPS activity is impacted by the herbicide. This is consistent with an NTSR mechanism being responsible for tolerance in *M. polymorpha*. Glyphosate treatment caused a higher fold increase in shikimate in *A. thaliana* compared to *M. polymorpha*, suggesting that there is more inhibition of EPSPS in *A. thaliana* than in *M. polymorpha*. This lower accumulation of shikimate in *M. polymorpha*, despite higher concentrations of glyphosate in the tissues, could indicate less inhibition of the shikimate pathway, a slower flow of metabolites through the shikimate pathway, or a possible induced change in the regulation of the shikimate pathway upon glyphosate treatment. This observation may also reflect MurA catalysing the reaction of EPSPS in *M. polymorpha* facilitating flux through the shikimate pathway. Furthermore, lower shikimate accumulation in *M. polymorpha* may directly confer glyphosate tolerance in *M. polymorpha* as shikimate to slightly reduces plant area in *A. thaliana* and not *M. polymorpha* (Section 5.4.3). This may not be a result of direct shikimate toxicity, but instead due to the tight regulation of the shikimate pathway impacted by shikimate accumulation.

Shikimate accumulation is likely not the primary cause of plant growth reduction in glyphosate-sensitive plants. Conversely, I demonstrated that co-treatment of *A. thaliana* plants with glyphosate and aromatic amino acids can prevent the lethal effects of glyphosate. This evidence is consistent with starvation of aromatic amino acids being the primary cause of plant death in glyphosate treated plants (Sammons and Gaines, 2014). However, when the pools of amino acids were measured using LC-MS/MS in *A. thaliana* and *M. polymorpha* following glyphosate treatment, contrary to my initial predictions, there was no decrease in the pools of the aromatic amino acids phenylalanine, tryptophan, and

tyrosine in *A. thaliana*. This suggests that glyphosate does not immediately inhibit the biosynthesis of aromatic amino acids in sensitive plants or secondary effects mask this inhibition. *M. polymorpha* displayed a significant increase in phenylalanine and tyrosine pools after 2 days of glyphosate treatment. This could be part of a response which works to maintain or even enhance the pools of aromatic amino acids following inhibition of the shikimate pathway. More pronounced changes in aromatic amino acid pools were observed after 7 days of glyphosate treatment. In *A. thaliana*, there was an increase in the quantity of all the aromatic amino acids, potentially due to protein catabolism as part of a stress response, or up-regulation of the shikimate pathway to compensate for its inhibition. After 7 days of glyphosate treatment, *M. polymorpha* exhibited a different response characterised by increased levels of phenylalanine and tyrosine, but a decrease in tryptophan. This difference may be attributed to the different biosynthetic pathway of tryptophan compared to the other aromatic amino acids and different regulation of this pathway in *M. polymorpha*.

Other amino acid pools also displayed significant differences in fold changes between the two species. For example, there was a substantial increase in proline levels in *A. thaliana* compared to *M. polymorpha*. Proline is known to accumulate in response to various environmental stresses, including herbicide treatment, suggesting a strong stress response elicited by glyphosate treatment in *A. thaliana* (Alia and Saradhi, 1991; Claussen, 2005). However, proline upregulation has also been associated with a response to oxidative stress, and so may suggest a role glyphosate plays in increasing reactive oxygen species (ROS) in sensitive plants (Ahsan et al., 2008; Alia, 2002). There is a large accumulation of arginine *M. polymorpha* compared to *A. thaliana*. This may indicate that glyphosate treatment also causes a stress response in *M. polymorpha*, since arginine can act as a precursor for the synthesis of polyamines that regulate cell growth, differentiation, and stress tolerance (Tiburcio et al., 2014). The distinct changes in amino acid pools in *A. thaliana* and *M. polymorpha* suggest that both undergo a stress response following glyphosate treatment, but that these may be distinct metabolic responses to glyphosate-induced stress.

Differences in the fold changes of quinic acid, chorismate and PEP were also measured with LC-MS/MS following glyphosate treatment in *A. thaliana* and *M. polymorpha*. These

differences are indicative of differential metabolic responses in the two species, which may contribute to, or be a result of, their contrasting levels of glyphosate tolerance. The increase in quinic acid levels in both *A. thaliana* and *M. polymorpha* suggests this is a conserved response to the inhibition of EPSPS. Quinic acid is an intermediate of the shikimate pathway and may accumulate due to inhibition of EPSPS or could have its biosynthesis increased through feedback from alternative metabolic pathways. The difference in quinic acid changes between the two species may represent variation in the ability to channel carbon and energy flow towards quinic acid synthesis. It has been postulated that quinic acid accumulation may contribute to glyphosate causing plant death (Orcaray et al., 2010), however, an accumulation of quinic acid in glyphosate-tolerant *M. polymorpha* is inconsistent with this. Chorismate is an immediate product of the shikimate pathway and exhibited a distinct response in *A. thaliana* and *M. polymorpha* after 2 days of glyphosate treatment. While *M. polymorpha* showed a decrease in chorismate levels, indicative of shikimate pathway inhibition, *A. thaliana* displayed an increase in chorismate, implying pathway upregulation or an alternative accumulation pathway (Lynch, 2022). After 7 days of glyphosate treatment, both species showed increased chorismate levels, indicating differential or staggered metabolic responses. Glyphosate treatment resulted in varying changes of PEP in *A. thaliana* and *M. polymorpha* after both 2 and 7 days of glyphosate treatment. *A. thaliana* exhibited increased PEP levels, while *M. polymorpha* displayed a decrease. In *M. polymorpha*, this decrease may indicate PEP utilisation in other pathways like the TCA cycle or glycolysis to redirect energy flow due to shikimate pathway inhibition or as a substrate for MurA. Conversely, *A. thaliana* may not efficiently redirect PEP, potentially leading to its accumulation post-glyphosate treatment, given its role as an EPSPS substrate. The effect of glyphosate on shikimate pathway intermediates in both *A. thaliana* and *M. polymorpha* suggest a broader metabolic response which included other pathways.

Using a non-targeted metabolic approach, substantial changes in the quantity of a variety of metabolites were detected in both *A. thaliana* and *M. polymorpha* following glyphosate treatment. In *A. thaliana*, most metabolites displayed an increase in abundance after both 2 and 7 days of glyphosate treatment. This suggests a general upregulation of metabolic pathways, potentially indicating a widespread stress response working to compensate for inhibition of the shikimate pathway. However, this trend may not be an evolved stress

response, but the result of the breaking down of various metabolic pathways, leading to the accumulation of various metabolites. In contrast, *M. polymorpha* had a more balanced metabolic response to glyphosate treatment, with close to equal numbers of metabolites increasing and decreasing in quantity following glyphosate treatment. The number of metabolites differentially regulated in *M. polymorpha* was also higher overall. This suggests a broader, more balanced, strategy utilised by *M. polymorpha* in response to glyphosate-induced stress, which may be more effective. Very few of the differentially regulated metabolites were common to both species, suggesting not only the magnitude, but also the type of response was entirely different. This likely reflects their distant evolutionary relationship, and since their last common ancestor, having evolved separate metabolic mechanisms for dealing with abiotic stress.

The differing metabolic responses to glyphosate in *A. thaliana* and *M. polymorpha* highlight the complexity of how glyphosate causes plant death. To understand more fully the metabolic pathways affected by glyphosate treatment, a pathway analysis was carried out. As anticipated, 2 days of glyphosate treatment significantly impacted the biosynthesis of aromatic amino acids in both *A. thaliana* and *M, polymorpha*, as glyphosate inhibits EPSPS, a crucial enzyme in this biosynthetic pathway. The observed impact on this pathway is consistent with the understood mechanism of glyphosate action (Duke, 2018). However, the analysis revealed the impact on aromatic amino acid biosynthesis was transient, as after 7 days of glyphosate treatment there was no longer a significant effect. This may indicate changes in the regulation of the shikimate pathway or the activation of alternative metabolic or catabolic routes to ensure sufficient supply of aromatic amino acids for essential processes. My analysis also identified significant and different impacts on other amino acid-related pathways, including alanine, aspartate, and glutamine metabolism, and arginine biosynthesis in *A. thaliana* and *M. polymorpha*. Therefore, glyphosate treatment induces species-specific effects on these pathways, potentially contributing to glyphosate tolerance in *M. polymorpha* through alterations in amino acid metabolism, affecting various cellular processes, including stress response and growth regulation.

Another pathway significantly affected by glyphosate treatment in both species was the flavone and flavanol biosynthesis pathway. Phenylalanine, a product of the shikimate

pathway, is a precursor to flavonoids, suggesting a potential link between glyphosate-induced changes in aromatic amino acid biosynthesis and alterations in flavonoid biosynthesis. Flavonoids play a crucial role in plant defence and signalling (Ghitti et al., 2022; Treutter, 2006). Therefore, altered regulation following glyphosate treatment could impact stress responses and defence mechanisms, which could contribute to glyphosate tolerance in *M. polymorpha*. Further targeted analysis of the flavone and flavanol biosynthesis pathway is needed to determine the impact of glyphosate treatment on this pathway and how this may relate to glyphosate mode of action.

My pathway analysis identified impact on the TCA cycle after 7 days of glyphosate treatment. The TCA cycle is a central metabolic pathway involved in energy production and was impacted in both *A. thaliana* and *M. polymorpha*. The TCA cycle intersects with the shikimate pathway, suggesting potential co-regulation and interconnected responses to glyphosate treatment. Glyphosate-induced alterations in the TCA cycle may have implications for energy metabolism and overall plant physiology, affecting processes such as respiration and carbon utilisation. Therefore, differential regulation of the TCA cycle following glyphosate treatment in *A. thaliana* and *M. polymorpha* may offer insights into the mechanism of glyphosate tolerance in *M. polymorpha*.

I employed a targeted metabolomics approach to resolve how the TCA cycle was impacted by glyphosate treatment in both *A. thaliana* and *M. polymorpha*. Changes in the TCA cycle can have far-reaching effects on plant physiology which may be linked to glyphosate tolerance in *M. polymorpha*. Consistent with this, after 2 days of glyphosate treatment there was a contrasting response in the TCA cycle in *M. polymorpha* and *A. thaliana*. In *A. thaliana* there was little to no change in the quantity of the measured TCA cycle intermediates, whereas in *M. polymorpha* I observed significant decreases in citrate, α-ketoglutarate, and malate. After 7 days of glyphosate treatment there were more substantial changes in the quantity of TCA cycle intermediates in both *A. thaliana* and *M. polymorpha*, indicating a prolonged metabolic response in this pathway to glyphosate exposure. In *A. thaliana*, all five intermediates measured accumulated after 7 days of treatment, suggesting an upregulation of the TCA cycle. This may be part of a stress response to glyphosate's inhibitory effects on the shikimate pathway, aiming to maintain cellular homeostasis and survival. In *M.*

polymorpha the changes in TCA cycle intermediates were mostly opposite to those observed in *A. thaliana*. While *cis*-aconitate showed significant accumulation, all other intermediates decreased in abundance. These reductions in TCA cycle intermediates in *M. polymorpha* may reflect a metabolic shift in response to glyphosate treatment, potentially associated with a mechanism to compensate for the inhibition of the shikimate pathway or to activate alternative pathways to maintain energy and biosynthetic demands. This may represent a unique metabolic adaptation to general stress or inhibition of the shikimate pathway, which is now being used to confer tolerance to glyphosate in *M. polymorpha*. The differing TCA cycle responses in *A. thaliana* and *M. polymorpha* highlight glyphosate's species-specific metabolic impacts. These differences suggest *M. polymorpha* may have evolved unique metabolic strategies to withstand environmental stress akin to glyphosate treatment, which allow essential metabolic functions and growth persistence in the presence of glyphosate.

Glycolysis is another fundamental metabolic pathway in plants but has fewer direct links to the shikimate pathway than the TCA cycle. Nevertheless, given its central role in energy metabolism, it is still expected that glyphosate treatment would impact glycolysis, particularly in the glyphosate sensitive *A. thaliana*. However, after 2 days of glyphosate treatment, there were no changes in the glycolytic intermediates measured in *A. thaliana*. This suggests that shortly after glyphosate treatment, there is no direct impact on glycolysis in the glyphosate sensitive species, consistent with the few direct links from glycolysis to the shikimate pathway. In *M. polymorpha* the two glycolytic intermediates, PEP and diphosphoglyceric acid, decreased significantly, consistent with a partial perturbation of glycolytic flux in the tolerant species. The differential response a short period after glyphosate treatment suggests *M. polymorpha* may initiate a mechanism to mitigate the effects of glyphosate treatment that involves changes in glycolysis. After 7 days of glyphosate treatment, glycolytic changes were more substantial in both species. However, there was still no change in the quantities of PEP, diphosphoglyceric acid, and pyruvate in *A. thaliana*, suggesting that glycolysis remains relatively stable despite prolonged glyphosate exposure. Conversely, in *M. polymorpha*, there was significant depletion in all glycolytic intermediates measured, except pyruvate, indicating a more substantial disruption or altered glycolysis regulation following longer exposure to glyphosate. The depletion of hexose-6-phosphate and DHAP in both species suggests similarities in the response, perhaps

with some delay in *A. thaliana*. However, the large differences in magnitude and direction of changes in other glycolytic intermediates may be attributed to differences in glyphosate sensitivity or differences in metabolic response to glyphosate treatment. The opposite impacts on PEP and diphosphoglyceric acid in *A. thaliana* and *M. polymorpha* suggest different glycolysis regulatory mechanisms in response to glyphosate treatment. These regulatory mechanisms could influence their differential responses to glyphosate and may be part of adaptive responses that determine their sensitivity or tolerance to glyphosate.

The metabolomic analysis of glyphosate-treated *A. thaliana* and *M. polymorpha* revealed extensive changes in the metabolome, even in the glyphosate-tolerant *M. polymorpha*. However, my RNA-seq analysis indicated minimal changes in response to glyphosate treatment in *M. polymorpha*, with only 26 differentially expressed genes identified. These findings suggest that transcriptome changes do not explain the metabolomic differences between *A. thaliana* and *M. polymorpha*, nor shed light on glyphosate tolerance mechanisms in *M. polymorpha*. *M. polymorpha* may have rapid post-transcriptional or translational regulatory mechanisms to influence metabolic responses to glyphosate treatment without extensive gene expression changes. Further RNA-seq analysis in *A. thaliana* is needed to explore this further.

To summarise, the metabolomic analysis of glyphosate-treated *A. thaliana* and *M. polymorpha* plants revealed extensive changes in the metabolome, even in the glyphosate-tolerant *M. polymorpha*, but minimal corresponding changes in gene expression in response to the herbicide. The lack of a strong correlation between metabolic and gene expression changes may imply the involvement of intricate regulatory mechanisms in *M. polymorpha*, possibly at the post-transcriptional or translational level, to regulate metabolic pathways to resist glyphosate's effects. Alternatively, in *M. polymorpha* there may be an inherent mechanism to regulate metabolic pathways in response to inhibition of the shikimate pathway that does not involve changes to gene expression. My results suggest that *M. polymorpha* may have evolved unique and efficient mechanisms to cope with induced stress without inducing significant shifts in gene expression. *M. polymorpha* may now be able to utilise these mechanisms to survive applications of glyphosate lethal to most plants. These mechanisms may be borne from the functional redundancy between EPSPS and MurA which

may have led to a change in the way the shikimate pathway is regulated in non-angiosperm streptophyte plants that encode a MurA enzyme.

5.6. Supplemetary data

Locus name	Gene function	Log2 Fold Change	Adjusted p-value
Mp4g15070	ribulose-bisphosphate carboxylase small chain	-0.593472729	5.61E-05
Mp6g09310	unknown	-1.861631566	5.61E-05
Mp4g15080	ribulose-bisphosphate carboxylase small chain	-0.558710234	9.35E-05
Mp4g02720	plastocyanin	-0.870822653	0.000170589
Mp5g10700	peroxidase	1.980236637	0.001756859
Mp2g26040	Microtubule-binding protein (translationally controlled tumor protein)	-0.72870613	0.006942753
Mp2g07830	ATPF0B, atpF, F-type H+-transporting ATPase subunit b	-1.499818817	0.01356178
Mp2g12030	unknown	-0.29917553	0.01442012
Mp4g10850	ribulose-bisphosphate carboxylase small chain	-0.286242228	0.01442012
Mp8g05670	unknown	0.92189728	0.01442012
Mp5g13860	GERMIN-LIKE PROTEIN SUBFAMILY 3 MEMBER 3	-0.886492345	0.016545274
Mp1g11470	RP-L35Ae, RPL35A, large subunit ribosomal protein L35Ae	-0.673354624	0.025629474
Mp2g08510	Predicted Ca2+-dependent phospholipid-binding protein	-1.554548984	0.025629474
Mp2g12020	unknown	-0.300076168	0.025629474
Mp5g15120	unknown	-1.254003331	0.025629474
Mp6g21460	Photosystem II reaction centre X protein (PsbX)	-0.492139487	0.025629474
Mp7g09180	light-harvesting complex II chlorophyll a/b binding protein 2	-0.334600861	0.025629474
Mp4g06030	ribulose-bisphosphate carboxylase small chain	-0.491333868	0.026815425
Mp8g10900	small subunit ribosomal protein S14e	-0.772311202	0.026815425
Mp4g16300	HIGH-AFFINITY NITRATE TRANSPORTER	-1.085495753	0.026940381
Mp4g23990	large subunit ribosomal protein L10e	-0.610932091	0.026940381
Mp1g20890	aquaporin TIP	-0.38720031	0.027492083
Mp1g16730	small subunit ribosomal protein S3e	-0.658298005	0.034435506
Mp1g14290	small subunit ribosomal protein S16e	-0.797866004	0.039397273
Mp4g06050	ribulose-bisphosphate carboxylase small chain	-0.360865126	0.042831341
Mp1g04920	large subunit ribosomal protein L13e	-0.937402668	0.046042897

Supplementary Table S5.1. All differentially expressed genes identified in *M. polymorpha* over a 48 h period

The differentially expressed genes identified from RNA-seq data when comparisons were made between many permutations of treatment and timepoint.

Chapter 6: General Discussion

Prolonged and intensive use of herbicides has led to the evolution of herbicide-resistant weed populations (Gaines et al., 2020; Powles and Yu, 2010). Herbicide resistance mechanisms can be categorised as target-site resistance (TSR) or non-target site resistance (NTSR). NTSR is particularly problematic as it can confer resistance to many different herbicides and is often under polygenic control (Délye et al., 2013). Four protein superfamilies – P450s, GSTs, UGTs and ABC transporters – have been most frequently implicated in NTSR. In chapter 2, I determined the phylogenetic relationships of ABC transporters and UGTs within the Archaeplastida lineage. I showed that ABC transporters were present in the last common ancestor of the Archaeplastida, which existed approximately 1.6 billion years ago, and major clades of ABC transporters in extant plants were already present then. By contrast, UGTs were only identified in members of the streptophyte lineage, suggesting either a loss of these proteins in the chlorophyte, rhodophyte and glaucophyte lineages or horizontal gene transfer of UGTs into a common ancestor of the streptophytes from a non-Archaeplastida eukaryote lineage. Most of the diversity of UGTs in extant plants evolved in the vascular plant lineage, likely accompanying an increase in the morphological complexity of plants following the colonisation of terrestrial habitats. Clade diversity of both ABC transporters and UGT enzymes in extant plants may contribute to their ability to confer herbicide resistance in populations of weeds. I conclude that both ABC transporters and UGTs are ancient protein families which have diversified and may have evolved developmental functions as plants began to occupy new terrestrial niches and are now being selected to confer NTSR in weeds. Such a process, wherein proteins evolved separate functions in plants but are now conferring NTSR, can be considered an example of exaptation.

Bryophytes are one of the two monophyletic groups of land plants, with the other being vascular plants (tracheophytes). Bryophytes comprise, liverworts, mosses, and hornworts. In chapter 3, I showed that the liverwort *Marchantia polymorpha* has inherent resistance (tolerance) to the herbicide glyphosate. Glyphosate is a widely and intensively used herbicide, and as a result, many weed species have evolved resistance to this herbicide (Heap and Duke, 2018; Heap, 2023). I hypothesised that *M. polymorpha* is naturally tolerant to glyphosate, a rare feature among angiosperms, and therefore potentially exhibits a novel form of tolerance. High levels of glyphosate resistance in weed plants are generally

associated with TSR (Gaines et al., 2020). However, despite predicting altered binding of glyphosate to MpEPSPS homology models, I showed the MpEPSPS enzyme was as sensitive to glyphosate as observed in other land plant EPSPS enzymes *in vitro*. Therefore, I conclude *M. polymorpha* is not tolerant to glyphosate because of a glyphosate insensitive EPSPS enzyme.

I hypothesised that EPSPS catalytic activity may influence glyphosate tolerance in *M. polymorpha*. To test this hypothesis, I conducted enzyme kinetic assays and found that the enzymatic efficiency of MpEPSPS is lower than EPSPSs from other land plants. Glyphosate inhibits the shikimate pathway, a complex and precisely regulated metabolic pathway which leads to the synthesis of aromatic amino acids in plants. It has been hypothesised that deregulation of the shikimate pathway upon glyphosate treatment could contribute to plant death (Siehl, 1997). I therefore hypothesised that the regulatory mechanisms of the shikimate pathway in *M. polymorpha* (explored in chapter 5), may be distinct from those in angiosperms. This distinction may be associated with the different kinetic properties of MpEPSPS, potentially contributing to tolerance. Using sequence alignments, homology modelling and molecular docking analyses, I identified a serine at position 500 in MpEPSPS, which I predicted to be responsible for low catalytic efficiency because it influences the structure of the active site of MpEPSPS compared to glyphosate sensitive EPSPS enzymes. However, through site-directed mutagenesis of this residue in transgenic *A. thaliana* plants overexpressing both the endogenous EPSPS and MpEPSPS, I was unable to show this residue alone could be attributed to the properties of MpEPSPS observed. This outcome may be the result of limitations in the homology modelling analysis and highlights the intricate relationship between sequence and catalytic function in EPSPS enzymes. This indicates that the serine at position 500 is likely not the sole cause of the reduced catalytic efficiency of MpEPSPS.

In chapter 4, I demonstrated that the peptidoglycan biosynthesis enzyme UDP-N-acetylglucosamine-3-enolpyruvyl transferase (MurA) confers glyphosate tolerance in *M. polymorpha*. I showed that EPSPS and MurA are structurally very similar, likely a result of both enzymes catalysing the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) to their respective substrates. Consistent with the role of MurA in conferring

glyphosate tolerance, I showed angiosperms do not encode a MurA enzyme, and two bryophytes – *Physcomitrium patens* and *Lunularia cruciata* – that encode MurA, are glyphosate tolerant. I confirmed the role of MurA in glyphosate tolerance in *M. polymorpha* by generating nine lines overexpressing MurA, five of which were significantly more resistant to glyphosate than wild-type. Furthermore, I generated seven *murA* loss-of-function mutants, all of which were significantly more sensitive to glyphosate compared to wild-type. Since MurA is involved in peptidoglycan biosynthesis, and peptidoglycan surrounds the chloroplasts of some plants (Hirano et al., 2016; Hofmann, 2016; Machida et al., 2006), loss-of-function of this enzyme led to a macrochloroplast phenotype in *M. polymorpha* plants. Glyphosate targets EPSPS, which is a chloroplast localised enzyme. To test if changes in chloroplast structure, due to a lack of peptidoglycan, and not MurA alone, contributed to observed changes in glyphosate resistance in *murA* mutants and MurA overexpressing plants, I generated five *murF* loss-of-function mutants. MurF is another essential enzyme in peptidoglycan biosynthesis. Furthermore, I treated wild-type *M. polymorpha* plants with peptidoglycan inhibiting herbicides. In both cases, macrochloroplasty was observed suggesting disruption of peptidoglycan biosynthesis, however, in neither case was the resistance to glyphosate of the plants altered. I therefore conclude that MpMurA contributes to glyphosate tolerance in *M. polymorpha* and does so independently of peptidoglycan biosynthesis.

I hypothesised that MurA may confer glyphosate tolerance in *M. polymorpha* by binding glyphosate (hypothesis 1), or by being catalytically active with the substrates of EPSPS (hypothesis 2). These hypotheses are not necessarily mutually exclusive. Hypothesis 1 suggests that MurA may confer glyphosate tolerance in *M. polymorpha* by binding to glyphosate, which in turn, would reduce the availability of glyphosate to inhibit EPSPS. Consequently, a larger proportion of EPSPS enzymes would remain active in the presence of glyphosate. Hypothesis 2 would confer glyphosate tolerance because MurA activity with the substrates of EPSPS would allow flux through the shikimate pathway even in the presence of glyphosate. This mechanism of tolerance would be analogous to target-site overexpression, the basis of herbicide resistance which has evolved in many weed species (Powles and Yu, 2010). In summary, these hypotheses represent testable mechanisms for how MurA may confer glyphosate tolerance in *M. polymorpha*.

To test hypothesis 1 – that MurA confers tolerance by binding glyphosate – I overexpressed various versions of MpMurA designed to be catalytically inactive but have varying portions of the active site intact. Furthermore, I carried out a micro-scale thermophoresis analysis on purified MpMurA enzymes to test for binding of glyphosate *in vitro*. With both techniques, I found no evidence that glyphosate binds to MurA. However, hypothesis 1 cannot be ruled out because the structural alterations introduced to the active site of MurA may have been too large to allow glyphosate binding, and MST measurements can be prone to false negative results (López-Méndez et al., 2021b, 2021a; Scheuermann et al., 2016).

To test hypothesis 2, I assayed the activity of purified MpMurA in the presence of the substrates of the reaction EPSPS catalyses – shikimate-3-phospate (S3P) and PEP. I detected the accumulation of phosphate colorimetrically and the other product of the reaction, 5-enolpyruvyl-shikimate-3-phosphate (EPSP), using LC-MS/MS. Detection of both of these products is consistent with MurA catalysing the transfer of the enolpyruvyl moiety from PEP to S3P. Consistent with hypothesis 2, this suggests that MpMurA has the capacity to facilitate flux through the shikimate pathway, and therefore contributes to glyphosate tolerance in *M. polymorpha* in this way. However, I also showed that the quantity of EPSP produced by MpMurA follows a dose-dependent relationship with glyphosate in a similar way to MpEPSPS, suggesting that both MpMurA and MpEPSPS are strongly inhibited by the herbicide. This finding is inconsistent with my observation that glyphosate does not bind to MpMurA and represents the only known example of glyphosate inhibiting a non-EPSPS enzyme. Furthermore, it also restricts the capacity for MpMurA to increase flux through the shikimate pathway in the presence of glyphosate, as both MurA and EPSPS are simultaneously inhibited. MurA likely then confers tolerance in *M. polymorpha* through a combination of hypothesis 1 and 2. Future research to determine the binding dynamics of glyphosate to MpMurA and the precise kinetic properties of MpMurA may help to resolve the exact mechanism of glyphosate tolerance in *M. polymorpha*.

The catalytic activity of MpMurA with the EPSPS substrates and its inhibition by glyphosate implies it is a target of glyphosate. The strength of TSR which evolves in weeds due to target overexpression is proportional to the level of expression of the target, and hence the

abundance of the target protein in the plant (Baerson et al., 2002; Gaines et al., 2010; Yanniccari et al., 2023). It is therefore formally possible that a high abundance of EPSPS and/or MurA enzymes in wild-type *M. polymorpha* plants may be responsible for glyphosate tolerance. Consistent with this, I showed that overexpression of MurA in *M. polymorpha* increased glyphosate resistance. Furthermore, expression of MurA in *M. polymorpha* could be sufficient to confer a higher level of glyphosate tolerance than plants that lack a MurA enzyme, such as angiosperms. In the future, it is therefore necessary to measure the relative abundance of these proteins in glyphosate-tolerant *M. polymorpha* compared to glyphosate-sensitive plants such as *A. thaliana* to determine if this factor contributes to glyphosate tolerance.

In chapter 5, I showed that the impacts of glyphosate treatment on the metabolome of the tolerant plant *M. polymorpha* differed from those on the glyphosate-sensitive plant *A. thaliana*. I showed that higher levels of shikimate accumulate in *A. thaliana* than *M. polymorpha*, suggesting greater relative inhibition of EPSPS activity in the angiosperm. I demonstrated that shikimate reduces growth rates of *A. thaliana* but not *M. polymorpha*, suggesting shikimate toxicity in the former species but not the latter may be responsible, at least in part, for the herbicidal effect of glyphosate in angiosperms. However, I propose that the primary cause of plant death in glyphosate-treated plants to be starvation of aromatic amino acids, as glyphosate lethality can be rescued with supplementation of aromatic amino acids in *A. thaliana*. However, using LC-MS/MS, I was unable to detect depletion of the pools of aromatic amino acids in either species. Instead, I observed higher levels of these amino acids and others following glyphosate treatment compared to untreated controls. This likely represents secondary effects of glyphosate treatment, which may mask the primary impacts causing plant death in *A. thaliana*.

I also conducted a non-targeted analysis of the metabolome of *A. thaliana* and *M. polymorpha,* which showed a widescale influence of glyphosate treatment on both the tolerant and sensitive species. However, these responses were often different in their magnitude and direction, suggesting species-specific regulatory responses to glyphosate treatment. These responses included energy metabolism pathways such as the TCA cycle and glycolysis, suggesting dysregulation of major pathways which may contribute to plant

death in *A. thaliana*, while compensatory responses may contribute to glyphosate tolerance in *M. polymorpha*. I also observed temporal changes in the response to glyphosate treatment. This is consistent with secondary impacts of treatment being observed, and therefore future work should focus on a shorter time period following treatment to better understand the primary effects of glyphosate treatment.

In conclusion, in this thesis I determined the phylogenetic relationships, within the Archaeplastida lineage, of the UGT and ABC transporter protein superfamilies, which have been widely reported to confer herbicide resistance in populations of weeds in the field. I report a rare case of natural glyphosate tolerance in the liverwort *M. polymorpha* and showed this is not the result of a tolerant form of the glyphosate target EPSPS, but a novel form of glyphosate tolerance involving the peptidoglycan biosynthesis enzyme MurA. I determined that this enzyme catalyses the production of EPSP and is inhibited by glyphosate. Therefore, MpMurA can be considered a novel glyphosate target. I showed the metabolomic responses to glyphosate treatment to differ between a glyphosate-sensitive and -tolerant species, suggesting metabolic regulation in *M. polymorpha* may contribute to tolerance. The findings of this thesis and future research based on them may be used to influence the efficiency of glyphosate use in the future.

Appendix

PLOS ONE

RESEARCH ARTICLE

ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation

Samuel Caygill[1,2], **Liam Dolan**[1,2]*

1 Gregor Mendel Institute, Vienna, Austria, **2** Department of Biology, University of Oxford, Oxford, United Kingdom

* liam.dolan@gmi.oeaw.ac.at

Check for updates

OPEN ACCESS

Citation: Caygill S, Dolan L (2023) ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation. PLoS ONE 18(9): e0287356. https://doi.org/10.1371/journal.pone.0287356

Editor: Yanbin Yin, University of Nebraska-Lincoln, UNITED STATES

Received: June 4, 2023

Accepted: August 25, 2023

Published: September 21, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0287356

Copyright: © 2023 Caygill, Dolan. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Abstract

ATP-binding cassette (ABC) transporters actively transport various substances across membranes, while uridine diphosphate (UDP) glycosyltransferases (UGTs) are proteins that catalyse the chemical modification of various organic compounds. Both of these protein superfamilies have been associated with conferring herbicide resistance in weeds. Little is known about the evolutionary history of these protein families in the Archaeplastida. To infer the evolutionary histories of these protein superfamilies, we compared protein sequences collected from 10 species which represent distinct lineages of the Archaeplastida–the lineage including glaucophyte algae, rhodophyte algae, chlorophyte algae and the streptophytes–and generated phylogenetic trees. We show that ABC transporters were present in the last common ancestor of the Archaeplastida which lived 1.6 billion years ago, and the major clades identified in extant plants were already present then. Conversely, we only identified UGTs in members of the streptophyte lineage, which suggests a loss of these proteins in earlier diverging Archaeplastida lineages or arrival of UGTs into a common ancestor of the streptophyte lineage through horizontal gene transfer from a non-Archaeplastida eukaryote lineage. We found that within the streptophyte lineage, most diversification of the UGT protein family occurred in the vascular lineage, with 17 of the 20 clades identified in extant plants present only in vascular plants. Based on our findings, we conclude that ABC transporters and UGTs are ancient protein families which diversified during Archaeplastida evolution, which may have evolved for developmental functions as plants began to occupy new environmental niches and are now being selected to confer resistance to a diverse range of herbicides in weeds.

Introduction

Populations of weeds are continually evolving resistance to herbicides [1]. Herbicide resistance mechanisms can be classified into 2 general groups: target site resistance (TSR) and non-target

Funding: This research was supported by a European Research Council (ERC) Advanced Grant De Novo-P (project number 787613) to LD from the European Commission. SC was supported by a British Biological Sciences Research Council (BBSRC) Scholarship through a doctoral training partnership (BB/M011224/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: L. D. is a co-founder of moa Technology Ltd. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

site resistance (NTSR). Target site resistance (TSR) occurs as a result of mutations in the gene coding for the herbicide's target protein, which causes reduced inhibition of the herbicide target. Non-target site resistance (NTSR) can occur due to a variety of mechanisms, but ultimately results in a reduction in the amount of herbicide reaching its target protein, or by reducing damage caused by the herbicide to the plant [2]. Some of the NTSR mechanisms reported include reduction in the uptake or translocation of the herbicide, metabolic detoxification of the herbicide and sequestration of the herbicide to a location where it is not active, such as the vacuole [3–6]. Diverse genetic changes affecting NTSR genes can lead to their encoded proteins contributing to herbicide resistance. For example, mutations leading to gene overexpression, the expression of a hyperactive form of the protein, loss-of-function mutations or mutations leading to changes in substrate specificity, can all result in NTSR [4, 7, 8]. The genetic basis of NTSR is poorly understood, but changes in the expression of both ATP-binding cassette (ABC) transporters and uridine diphosphate (UDP) glycosyltransferases (UGTs) have been shown to confer resistance to herbicides in weeds [3–5].

ABC transporters are one of the largest protein families in plants and are conserved across all domains of life [9]. ABC transporter proteins use the energy of ATP hydrolysis to drive active transport of substances across membranes [10, 11]. In plants, ABC transporters transport nutrients such as sugars, amino acids and metal ions; secondary metabolites such as flavonoids, terpenoids and alkaloids; phytohormones such as auxins, cytokinins and gibberellins; proteins and peptides; and xenobiotics [12–14]. Plant ABC transporters are also involved in the translocation of toxic compounds, such as herbicides, into the vacuole and extracellular spaces, where they are no longer toxic [15, 16]. Numerous ABC transporters have been identified to compartmentalise herbicides in this way, as part of a detoxification process that has evolved in some weeds with NTSR to at least one herbicide.

Most ABC transporters are membrane bound and consist of four subunits–two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs) (S1 Fig) [17]. Generally, the substrate binds to the TMDs of the transporter on one side of the membrane. Binding of ATP to the NBDs then causes a conformational change in the transporter, closing the TMDs around the substrate. Subsequently, the hydrolysis of ATP causes another conformational change in the transporter that opens up the TMDs on the opposite side of the membrane, releasing the substrate [10, 18]. It is with this mechanism that ABC transporters are able to transport herbicides as part of a herbicide-resistance-conferring detoxification process.

Agricultural fields have been treated with herbicides since the 1940s, exposing plants to chemical diversity that they would not have previously encountered. Some ABC transporters catalyse the transport of a wide variety of synthetic compounds across membranes that they would not have encountered in their evolutionary history. ABC transporters are one of the most diverse protein families in plants, exhibiting significant variation in sequence, structure, and function. In plants, ABC transporters have been classified into 8 subfamilies (ABCA, ABCB, ABCC, ABCD, ABCE, ABCF, ABCG and ABCI) [10, 11, 19, 20]. Previous phylogenetic analyses in plants have shown all ABC transporter subfamilies, excluding ABCI, constitute monophyletic clades [10, 11, 19]. However, previous phylogenetic analyses of these protein families focused primarily on angiosperms. To gain a more comprehensive understanding of the evolution of this protein family, here we carry out a phylogenetic analysis of ABC transporters from across the Archaeplastida.

Uridine diphosphate (UDP) glycosyltransferases (UGTs) are part of the largest family (Family 1) of the glycosyltransferase superfamily, and this subset of UGTs have been widely associated with herbicide resistance [21, 22]. UGTs are present in bacteria, fungi, plants, and animals and catalyse the transfer of glycosyl residues from activated nucleotide sugars, such as uridine diphosphate glucose, to acceptor molecules [23]. Acceptor molecules utilised by UGTs

include flavonoids, phenolic acids, plant hormones such as IAA, xenobiotics, and other small molecules such as alkaloids, terpenoids and cytokinins [24]. While UGTs often transfer glycosyl residues to various endogenous compounds, a variety of exogenous xenobiotics, including herbicides, can also act as the acceptor molecules of UGTs. The addition of a glycosyl residue to a herbicide has the general effect of making it more hydrophilic and less toxic. Therefore, mutations in genes coding for UGTs that enable them to glycosylate herbicides, or mutations that cause overexpression of UGTs that can already glycosylate herbicides, can contribute to NTSR in weeds.

UGTs are classified by the presence of a C-terminal consensus sequence, the UDP-glycosyltransferase signature, known as the 'PSPG' motif (Plant Secondary Product Glycosyltransferase). The PSPG motif is made up of 44-amino acids and is thought to be the site of binding of the UDP moiety of the sugar molecule to the enzymes [25–27]. On the other hand, the N-terminal regions of UGTs is the domain which is involved in the recognition and binding of a diverse array of acceptor molecules. Therefore, the N-terminal domain is generally more variable than the C-terminus [28]. UGTs catalyse O-, S- and N-linkages and can also form different linkages with the same acceptor molecule [29]. For example, *A*. *thaliana* UGT72B1 has been reported to form an N-glucosidic bond with 3, 4-dichloroaniline and an O-glucosidic bond with 3,4-dihydroxybenzoic acid [30]. Previous phylogenetic analyses of plant UGTs has found this protein family to diverge into at least 20 distinct monophyletic groups [25, 26]. However, the taxon sampling of these studies has been limited to angiosperms, and so conclusions cannot be made regarding the evolutionary history of this protein family in the Archaeplastida lineage.

Here we report the phylogenetic relationships, of both the ATP-binding cassette transporter and uridine diphosphate glycosyltransferase protein superfamilies within the Archaeplastida lineage. We discovered that the current clade diversity of the ABC transporter superfamily existed in the last common ancestor of the Archaeplastida. This indicates that the ABC transporters diversified into the extant sub-families before the evolution of the last common ancestor of the Archaeplastida, extant 1.6 billion years ago [31–33]. By contrast UGT proteins are only present in the streptophyte lineage among the Archaeplastida and diversified in the vascular plant lineages. On the other hand, we show that UGTs are present in the streptophyte lineage of the Archaeplastida but absent from non-streptophyte lineages. This suggests either that this protein family, which exists in other eukaryotic lineages, was lost from the glaucophyte, rhodophyte and chlorophyte lineages, or that these proteins originated through horizontal gene transfer into the last common ancestor of the streptophytes, extant approximately 1.1–1.2 billion years ago in the Proterozoic era [34–39]. Furthermore, we show that the UGT family subsequently diversified throughout streptophyte evolution, mostly in the vascular lineage. This analysis suggests that gene superfamilies which evolved in the Proterozoic era, are now undergoing natural selection caused by herbicide application which is leading to the evolution of herbicide resistant weeds in agricultural systems.

Materials and methods

Data resources

Amino acid sequences of proteins from *A*. *thaliana* were collected from TAIR10 [40] (https://www.arabidopsis.org/). Sequences from *O*. *sativa* were retrieved from the rice genome annotation project [41]. Sequences from the liverwort *M*. *polymorpha* were collected from MarpolBase (http://marchantia.info/). Sequences from the hornwort *Anthoceros agrestis* were collected from [42] (https://www.hornworts.uzh.ch/en/download.html). Sequences from the streptophyte alga *Klebsormidium nitens* were obtained from the *K*. *nitens* genpme webpage

[43]. (http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/). Sequences from the lycophyte *Selaginella moellendorffii*, the moss *Physcomitrium patens* and the chlorophyte alga *Chlamydomona reinhardtii* were retrieved from Phytozome 12 [44] (https://phytozome.jgi.doe.gov/pz/portal.html). Protein sequences from the rhodophyte alga *Cyanidioschyzon merolae* were collected from the *C. merolae* genome webpage [45] (https://www.genome.jp/kegg-bin/show_organism?org=cme). Sequences from the glaucophyte alga *Cyanophora paradoxa* were collected from PhycoCosm [46] (https://phycocosm.jgi.doe.gov/Cyapar1/Cyapar1.home.html).

Sequence collection

Family 1 UGT and ABC transporter protein sequences from *A. thaliana* and *O. sativa* [40, 41], were used to perform BLASTP searches against the predicted proteomes of *S. moellendorffii*, *M. polymorpha*, *A. agrestis*, *P. patens*, *K. nitens*, *C. reinhardtii* and *C. merolae*. A minimum E-value threshold of $1e^{-10}$ was employed. The initial list of sequences for each species collected in this way was then used to 'self-blast' the predicted proteomes of that species in order to retrieve sequences belonging to species-specific Family 1 UGT and ABC transporter sub-families. Using InterProScan 92.0 [47], each Family 1 UGT sequence was checked for the UDP-glycosyltransferase domain (SSF53756), while each ABC transporter was checked for the presence of the ABC transporter pfam domain (PF12848). All sequences that did not contain the searched-for domains were removed before further analysis.

Sequence alignment

Collected amino acid sequences were aligned using the programme MAFFT [48] and its FFT-NS-2 algorithm. Alignments were then visualised using Unipro ugene v. 45.1 [49]. Sequences which could then be seen to be lacking important functional residues were then removed. Sequences were trimmed manually, with large non-conserved regions being removed, leaving only mostly conserved regions without gaps.

Phylogenetic analysis

The final trimmed amino acid sequence alignments were subsequently used in a maximum-likelihood phylogenetic analysis using the software PHyML 3.0 [50]. The parameters used were an estimated gamma distribution, the LG+G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio test (aLRT). The resulting phylogenies were visualised in FigTree v1.4.4 [51] and annotated in Inkscape v1.0.2 [52].

Results

960 ABC transporter and 552 UGT sequences were identified in the genomes of 10 species of Archaeplastida

We first set out to investigate the evolution of ATP-binding cassette (ABC) transporter and uridine diphosphate (UDP) glycosltransferase (UGT) gene families in the Archaeplastida lineage. The Archaeplastida is the lineage which includes glaucophyte algae, rhodophyte algae, chlorophyte algae and the streptophytes. ABC-transporter protein-encoding sequences were identified from 10 selected species (Table 1), representing distinct Archaeplastida lineages. UGT sequences were identified in the genomes of 6 of these 10 Archaeplastda species. There were 60 ABCs and 3 UGTs in the genome of the streptophyte alga *Klebsormidium nitens*; 100 ABCs and 50 UGTs in the liverwort *Marchantia polymorpha*; 85 ABCs and 15 UGTs in the moss *Physcomitrium patens*; 166 ABCs and 163 UGTs in the lycophyte *Selaginella*

Table 1. List of members of Archaeplastida used in the analysis.

Species	Classification	Genome (Mb)	Protein-coding genes	ABCs	ABCs (% PCG)	UGTs	UGTs (% PCG)	References
Arabidopsis thaliana	Angiosperm eudicot	135	25,498	120	0.471	121	0.475	[53]
Oryza sativa	Angiosperm monocot	321	35,681	124	0.339	200	0.561	[54]
Selaginella moellendorffii	Lycophyte	212.6	22,285	166	0.745	163	0.731	[55]
Anthoceros agrestis	Hornwort	133	24,700	130	0.526	0	0	[42]
Physcomitrella patens	Moss	480	35,938	85	0.237	15	0.042	[56]
Marchantia polymorpha	Liverwort	225.8	19,138	100	0.523	50	0.261	[57]
Klebsormidium nitens	Streptophyte alga	117.1	16,215	60	0.370	3	0.0185	[43]
Chlamydomonas reinhardtii	Chlorophyte alga	120	15,143	47	0.310	0	0	[58]
Cyanidioschyzon merolae	Rhodophyte alga	16.5	5,331	17	0.319	0	0	[45]
Cyanophora paradoxa	Glaucophyte alga	98.4	25,518	111	0.435	0	0	[46]

Including their phylum (Classification), genome size (Genome), total number of protein-coding genes (Protein-coding genes), total number of ABC transporter proteins (ABCs), ABC transporter proteins as a percentage of total protein coding genes (ABCs % PCG), total number of UDP-glycosyltransferase proteins (UGTs), UDP-glycosyltransferase proteins as a percentage of total protein coding genes (UGTs % PCG) and the reference for each genome sequence.

https://doi.org/10.1371/journal.pone.0287356.t001

moellendorffii; 124 ABCs and 200 UGTs in the monocot *Oryza sativa*; and 120 ABCs and 121 UGTs in the eudicot *Arabidopsis thaliana*. UGT sequences were not recovered from the genomes of the glaucophyte algae *Cyanophora paradoxa*, the rhodophyte alga *Cyanidioschyzon merolae*, the chlorophyte alga *Chlamydomonas reinhardtii*, or the hornwort *Anthoceros agrestis*. While no UGT encoding genes were identified in these species, 111, 17, 47, and 130 ABC transporter coding sequences were identified in these species respectively. Taken together these data indicate that the ABC transporters were present in the genome of the last common ancestor of the Archaeplastida, while UGT encoding genes have been present in the Archaeplastida lineage since the divergence of the streptophytes from the last common ancestor shared with the chlorophytes.

Plant ABC transporters are ancient, and all 12 groups existed in the last common ancestor of the Archaeplastida

To investigate the evolutionary history of the ABC transporter protein family, we generated a multiple sequence alignment of 960 ABC transporter protein sequences, which we used to construct a Maximum Likelihood tree (Fig 1). Alignments were trimmed manually to remove large gaps and only a nucleotide-binding domain (NBD) of each sequence was retained.

The topology of the phylogenetic tree shows that plant ABC transporters diverged into 6 previously identified, monophyletic clades. Each of these monophyletic clades has been classified by a single sub-family of ABC-transporters, with each subfamily designated a letter. For example, the monophyletic clade comprising of ABCA proteins is designated clade A. However, ABC transporter sub-families ABCE and ABCF are sister clades, and together form a single monophyletic clade, designated clade E/F. Our analysis revealed the previously defined monophyletic clades A, B, C, D, E/F, G (Fig 1). We also identified 6 previously unreported monophyletic clades that comprise sequences from a single species (*C. paradoxa*–Cp) or two species (*C. paradoxa* and *C. merolae*–CpCm)–Cp1, Cp2, Cp3, Cp4, Cp5, and CpCm1 (Fig 1). ABCI proteins are the only previously reported sub-family that do not form a monophyletic clade. We found ABCI proteins to comprise two distantly related monophyletic groups (and therefore constitute a polyphyletic group) (Fig 1). Sequences from the 4 clades B, C, E/F and G are encoded in the genomes of all 10 species of Archaeplastida used in this analysis. This suggests that these four clades were present in the last common ancestor of the Archaeplastida (Fig 2).

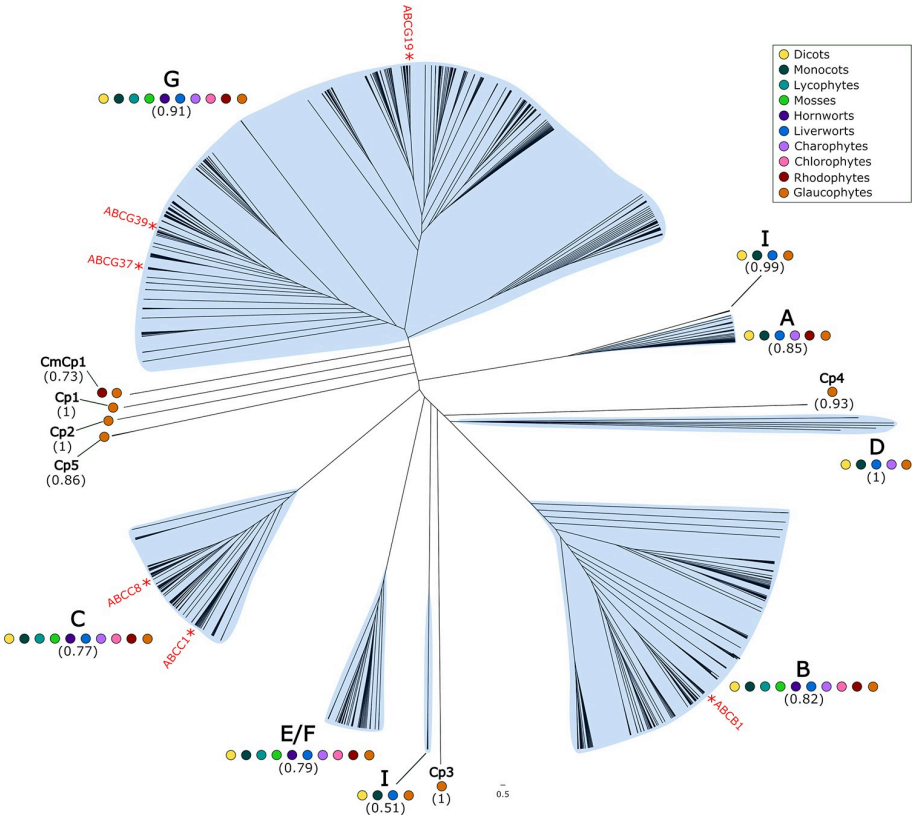


Fig 1. Phylogenetic analysis of ABC transporter protein sequences in the Archaeplastida. Unrooted cladogram of a maximum likelihood (ML) analysis of Archaeplastida ABC transporter proteins conducted by PhyML 3.0 [50] using an estimated gamma distribution parameter, the LG_G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio (aLRT) test. Support values from the aLRT test for each clade are presented in parentheses. Protein sequences were aligned using MAFFT with the L-INS-I algorithm. ABC transporter classes are indicated by blue highlighting and letters. Coloured dots indicate the presence of sequences from different species in each clade. *A. thaliana* (yellow); *O. sativa* (dark green); *S. moellendorffii* (light blue); *P. patens* (light green); *A. agrestis* (dark purple); *M. polymorpha* (blue); *K. nitens* (light purple); *C. reinhardtii* (pink); *C. merolae* (maroon); *C. paradoxa* (orange). Confirmed NTSR genes or their *A. thaliana* homolog reported in Table 3 are indicated.

https://doi.org/10.1371/journal.pone.0287356.g001

The presence of an ABCA protein sequences in the rhodophyte *C. merolae* and the glaucophyte *C. paradoxa* suggests clade A is ancient and was present in the last common ancestor of the Archaeplastida. However, ABCA proteins are missing from the genomes of 4 species–*C. reinhardtii*, *A. agrestis*, *P. patens* and *S. moellendorffii*. The absence of ABCA proteins from these lineages likely represent independent gene loss events, following the divergence of these lineages from other land plants. There are no ABCA proteins encoded in the genome of the chlorophyte alga *C. reinhardtii*, but ABCA proteins are encoded by the streptophyte alga *K. nitens*. This suggests a loss of these proteins in the chlorophyte algae lineage sometime after its divergence from the last common ancestor of the chlorophytes and streptophytes. ABCA proteins are present in the genome of the liverwort *M. polymorpha* but absent from the other two bryophyte species included in the analysis–the hornwort *A. agrestis* and the moss *P. patens*. If hornworts are sister to the mosses and liverworts [42, 59], this topology suggests ABCA proteins were lost from the hornwort lineage following its divergence from the moss and liverwort lineages. ABCA proteins were then likely subsequently lost in the moss lineage following its divergence from the liverworts. ABCA proteins were found to be absent from the genome of *S. moellendorffii* but present in the other 2 vascular

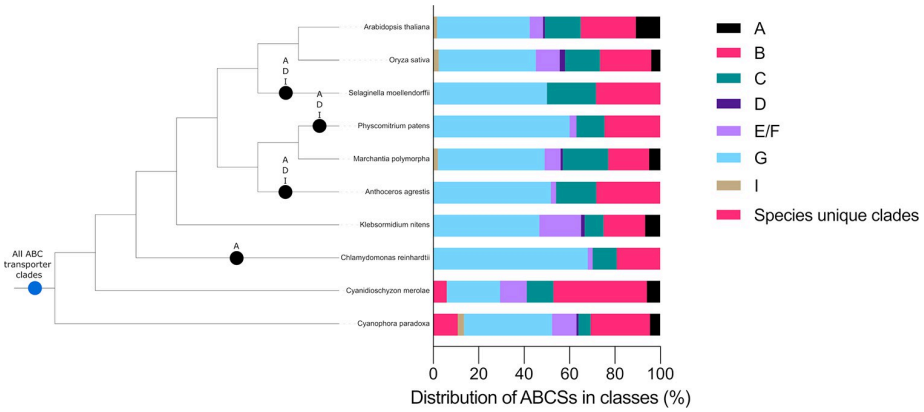


Fig 2. ABC transporters clades were present in the last common ancestor of the Archaeplastida. Cladogram of Archaeplastida showing ABC transporter clade origins and losses during plant evolution. Blue circles represent first appearances of a clade, black circles represent the loss of a clade in a particular lineage. Also shown are the relative distributions of ABC transporters into their respective clades for each species included in our analysis, where each distinct coloured block represents a separate ABC transporter protein clade.

https://doi.org/10.1371/journal.pone.0287356.g002

plants included in the analysis–*A*. *thaliana* and *O*. *sativa*. This indicates a loss of ABCA proteins in the lycophyte lineage following its divergence from the angiosperm lineage (Fig 2). The absence of ABCA proteins in some lineages within the Archaeplastida is likely due to ABCA proteins being secondarily lost in these lineages, and hence ABCA proteins were likely already present in the last common ancestor of the Archaeplastida.

ABCD proteins were present in 5 lineages of the Archaeplastida, the dicots, monocots, liverworts, streptophyte algae and glaucophyte algae. The presence of ABCD proteins in the genome of the glaucophyte *C*. *paradoxa* suggest that ABCD proteins had already diverged from other ABC transporters in the last common ancestor of the Archaeplastida. However, ABCD sequences were not found in *C*. *merolae* or *C*. *reinhardtii*, suggesting loss of the ABCD subfamily in the rhodophyte and chlorophyte algal lineages. ABCD proteins were missing from the genomes of *A*. *agrestis*, *P*. *patens* and *S*. *moellendorffii*, as observed for ABCA proteins. Their absence from these lineages suggests recent gene losses in the hornworts, mosses, and lycophytes respectively. Angiosperm genomes were found to encode ABCD proteins, however, the other vascular plant included in the analysis, *S*. *moellendorffii*, does not encode any ABCD proteins, suggesting that this clade was lost in this lineage following the divergence of the lycophyte lineage from the other vascular plants (Fig 2). ABCD proteins were therefore likely present in the last common ancestor of the Archaeplastida lineage, but were subsequently lost in the rhodophyte, chlorophyte, hornwort, moss and lycophyte lineages.

Many putative ABCI encoding genes were identified in the genomes of the plant species in this analysis. However, uniquely amongst ABC transporter proteins, each subunit of an ABCI protein is encoded by a separate gene. Therefore, many ABCI protein sequences identified were removed from the alignments generated as they did not align with the NBD of other ABC-transporters. In the phylogenetic analysis of the retained sequences, 2 monophyletic ABCI clades were defined. One was sister to the monophyletic ABCA protein clade, while the other is sister to a monophyletic clade that includes ABCD and ABCB proteins. Each of the ABCI clades comprise sequences from *A*. *thaliana*, *O*. *sativa*, *M*. *polymorpha* and *C*. *paradoxa*. ABCI proteins-encoding genes in the genome of *C*. *paradoxa* and the 3 previously mentioned land plant species is evidence that they were present in the last common ancestor of the Archaeplastida. However, absence of ABCI protein-encoding genes in the genomes of *C*.

merolae, *C. reinhardtii* and *K. nitens* suggests that ABCI protein-encoding genes have been lost in these algal lineages. Furthermore, the tree topology shows that losses of ABCI proteins have occurred in the land plants. The absence of ABCI proteins-encoding genes in *A. agrestis*, *P. patens* and *S. moellendorffii* suggest a loss of ABCI proteins in the hornwort, moss and lycophyte lineages (Fig 2). However, since many putative ABCI proteins were removed from the analysis, we cannot be entirely certain that members of this subfamily have been entirely lost from these lineages mentioned. The evolution of the ABCI gene clade may be more difficult to resolve compared to other ABC transporter clades, due to their unique property of individual genes encoding each protein subunit.

Taken together the data show that ABC transporters are conserved among the Archaeplastida. There are sequences from the 10 sampled species from the Archaeplastida that constitute four previously identified monophyletic clades, B, C, E/F and G (Fig 1). This suggests that ABC transporters had diverged into these four subfamilies before the divergence of the Archaeplastida lineages and were therefore present in the last common ancestor of all species included in the analysis. We also identified 6 other ancient monophyletic clades of ABC transporter which constituted sequences that were present in either glaucophyte, or both glaucophyte and rhodophyte algae (Cp1, Cp2, Cp3, Cp4, Cp5 and CpCm1). ABCA, ABCD and ABCI proteins are not present in the genomes of every species of the Archaeplastida analysed here. Our phylogenetic analysis revealed that although ABCA proteins were present in the last common ancestor of the Archaeplastida, members of this clade were subsequently lost in the chlorophyte, hornwort, moss and lycophyte lineages. We show that ABCD proteins also had an ancient origin, likely already having diverged in the last common ancestor of the Archaeplastida. ABCD proteins were subsequently lost in the rhodophyte, chlorophyte, hornwort, moss and lycophyte lineages. The ABCI proteins included in our analysis comprise 2, distantly related clades. Conservation of ABCI proteins included in our analysis was limited, however, their presence in *C. paradoxa* suggests an ancient origin of ABCI proteins. No ABCI protein-encoding sequences were found in the rhodophyte, chlorophyte, streptophyte algae, hornwort, moss or lycophyte lineages, which is consistent with the loss of genes encoding these proteins from these lineages. This suggests that ABCI proteins are ancient but have been lost from many lineages during Archaeplastida evolution.

Plant UGTs diversified recently, 14/20 identified groups are only present in Angiosperms

There are currently 45 known families of glycosyltransferases (GTs) in plants [60]. Uridine diphosphate glycosyltransferases (UGTs) are the largest glycosyltransferase family and are the only glycosyltransferase family to have been reported to contribute to NTSR. They comprise a monophyletic clade, known as Family 1 GTs. Family 28 GTs are a distinct glycosyltransferase protein family that are related to UGTs [23, 61–63]. Family 28 GTs have not been reported to contribute to NTSR. Family 28 GTs were included as an outgroup for our phylogenetic analysis. To investigate the evolutionary history of UGTs in Archaeplastida, phylogenetic trees were constructed using maximum likelihood statistics (Fig 3). This analysis showed that UGTs are only present in the streptophyte lineage of the Archaeplastida. UGTs are present in other eukaryotes, suggesting they were either lost in the glaucophyte, rhodophyte and chlorophyte lineages, or were introduced into the last common ancestor of the streptophyte lineage by horizontal gene transfer. Evidence of horizontal gene transfer of a glycosyltransferase gene has been previously reported from a Verrucomicrobium into the green alga *Picochlorum* SENEW3 [64]. Our phylogenetic analysis also shows UGT sequences comprise 20 major monophyletic groups, 16 of which were defined in previous analyses [25, 26].

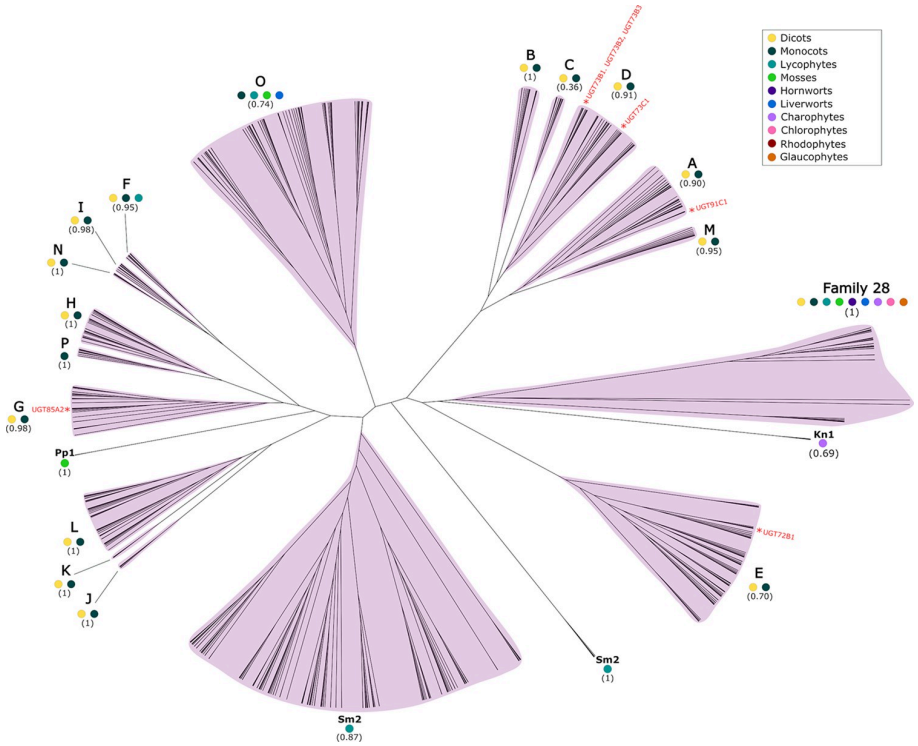


Fig 3. Phylogenetic analysis of UGT protein sequences in the Archaeplastida. Unrooted cladogram of a maximum likelihood (ML) analysis of Archaeplastida UGT proteins conducted by PhyML 3.0 [50] using an estimated gamma distribution parameter, the LG_G+F model of amino acid substitution, and a Chi^2-based approximate likelihood ratio (aLRT) test. Support values from the aLRT test for each clade are presented in parentheses. Protein sequences were aligned using MAFFT with the L-INS-I algorithm. UGT classes are indicated by pink highlighting and letters. Coloured dots indicate the presence of sequences from different species in each group. *A. thaliana* (yellow); *O. sativa* (dark green); *S. moellendorffii* (light blue); *P. patens* (light green); *A. agrestis* (dark purple); *M. polymorpha* (blue); *K. nitens* (light purple); *C. reinhardtii* (pink); *C. merolae* (maroon); *C. paradoxa* (orange). Confirmed NTSR genes or their *A. thaliana* homolog reported in Table 5 are indicated.

https://doi.org/10.1371/journal.pone.0287356.g003

34 Family 28 GT sequences were identified in the genomes of the 10 Archaeplastida species, and these were included in the phylogenetic analysis. Family 28 GT sequences were found in 9 of the 10 Archaeplastida species but are missing from the rhodophyte alga *C. merolae*. This is consistent with Family 28 GTs being present in the last common ancestor of the Archaeplastida but subsequently lost in the rhodophyte algae lineage (Fig 4). This likely represents separate origins of Family 28 GTs and UGTs in the Archaeplastida, the latter of which is only present in the streptophyte lineage.

The UGT protein family has gone through many diversification events since the divergence of Family 28 GTs and UGTs from a last common ancestor. We identified 20 distinct monophyletic groups of UGTs, 16 of which had been previously identified. Previously reported groups are named alphabetically. Each of the 4 new groups identified in this analysis are comprised of UGT sequences from a single species and are named according to that species. The streptophyte algae *K. nitens* is the earliest diverging member of the Archaeplastida found with UGT protein-encoding genes in its genome. Three UGTs were identified in *K. nitens* and these formed a single monophyletic group (Kn1), sister to all other UGTs identified in the analysis. This is consistent with UGTs being present in the last common ancestor of the streptophytes since *K. nitens* is the earliest diverging member of the streptophyte lineage included

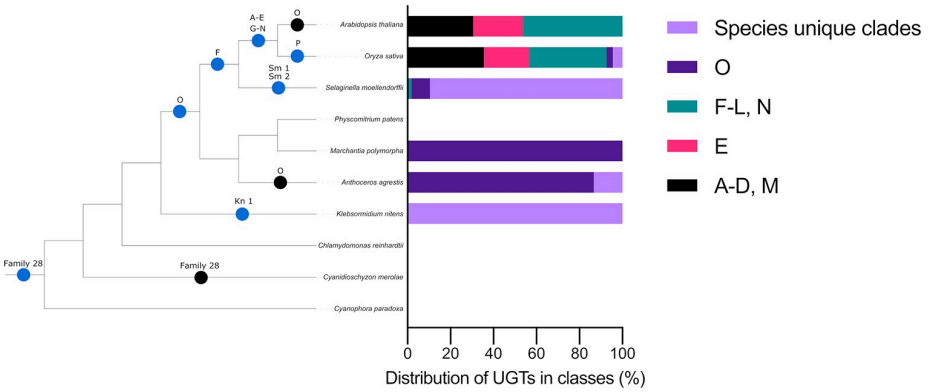


Fig 4. UGTs are present in streptophytes and have diversified into many clades in the vascular lineage Angiosperms. Cladogram of Archaeplastida phylogeny showing UGT clade origins and losses during plant evolution. The origins and losses of Family 28 glycosyltransferases are also shown. Blue circles represent first appearances of a clade, black circles represent the absence of a clade in a particular lineage. Also shown are the relative distributions of UGTs into their respective clades for each species included in our analysis, where each distinct coloured block represents either a separate UGT clade or a group of clades.

https://doi.org/10.1371/journal.pone.0287356.g004

in this analysis. The group O clade comprises 83 UGTs from 4 species of the Archaeplastida–*O. sativa*, *S. moellendorffii*, *P. patens* and *M. polymorpha*. Conservation of these sequences in these 4 land plant species and their absence from the genomes of the streptophyte algae indicates that this clade originated after the divergence of land plants from the streptophyte algae. Group O UGTs sequences are not present in the genomes of *A. thaliana* and *A. agrestis*, suggesting a loss of genes in the eudicot lineage after the divergence of *O. sativa* and *A. thaliana* from their last common ancestor, and in the hornwort lineage respectively. UGTs sequences from *O. sativa*, *S. moellendorffii*, *P. patens*, *M. polymorpha* and also *A. thaliana* constitute a large monophyletic clade consisting of UGT protein from groups F, G, H, I, J, K, L, N, P and Pp1. This suggests that this clade originated at approximately the same stage as group O UGTs, and later diverged into each of the 10 groups [25, 26]. This evidence suggests few divergence events of UGT groups occurred during early streptophyte evolution and then more prolific diversification in more recently diverging lineages.

13 of the 20 UGT groups–A, B, C, D, E, G, H, I, J, K, L, M, N–are only present in the genomes of *A. thaliana* and *O. sativa*. This is consistent with these groups being present in the last common ancestor of the angiosperms (Fig 4). Since these sequences are not present in the genome of *S. moellendorffii*, it suggests that they originated after the divergence of the seed plant lineage from the lycophytes. There are group F UGT sequences in each of the 3 vascular plant species included in the analysis, suggesting this group originated in the last common ancestor of the vascular plant lineage. Group F is sister to groups N and I, which comprise only angiosperm sequences. This suggests groups N and I diverged from a common ancestor which they shared with group F. These data show most UGT diversification in extant taxa occurred after the divergence of the seed plants from the lycophytes.

Five groups were defined by sequences from a singles species in our analysis. Group P comprises UGT sequences from only the monocot angiosperm *O. sativa*. There are sequences from only the lycophyte *S. moellendorffii* in groups Sm1 and Sm2, sequences from only the streptophyte algae *K. nitens* in the group Kn1, and there are sequences only from the moss *P. patens* in group Pp1. The formation of these monospecific clades is consistent with recent origins of these sequences but could also be an artifact of poor taxon sampling in this analysis. Group P is only found in *O. sativa* and not the other angiosperm *A. thaliana*, suggesting that group P

originated after the divergence of the monocots and the eudicots or was lost in the dicot lineage. Group Sm1 is the largest clade identified in our analysis and comprises 141 UGT sequences. Sm1 sequences were only found in the *S. moellendorffii* genome. This is consistent with this group originating in the lycophyte lineage, following the divergence of lycophytes from other vascular plants. The large number of sequences in this group suggest a prolific expansion of the Sm1 UGTs in the lycophyte lineage. Sequences from group Sm2 were also identified only in *S. moellendorffii*, indicating this group also originated in the lycophyte lineage. However, with 6 sequences in this group, Sm2 did not go through the same extensive expansion as Sm1. Streptophyte algae sequences from *K. nitens* form a monophyletic clade (Kn1) which is sister to all other UGTs. Kn1 likely represents the earliest diverging UGTs identified in this analysis. Two moss sequences form a monophyletic clade (Pp1). Pp1 is sister to a larger monophyletic clade consisting of groups F, G, H, I, P and N. This suggests Pp1 genes share a common ancestor with sequences from groups F, G, H, I, P and N which that existed before the divergence of bryophytes from other land plants. However, the absence of Pp1 genes in *M. polymorpha* and *A. agrestis* suggest that sequences from this group were lost in these lineages. The origins and losses of these groups in different Archaeplastida lineages are summarised in Fig 4.

In summary, we show UGT sequences are present in streptophyte algae but not in glaucophytes, rhodophytes or chlorophytes; UGTs are encoded in the *K. nitens* genome and in land plants except the hornworts where we propose they were secondarily lost. This indicates that UGTs first arose after the divergence of the streptophyte lineage and continued to diversify following the colonisation of land.

Six ABC-transporter sub-families have been associated with non-target site resistance to herbicides

ABC transporters have been implicated in non-target site herbicide resistance (NTSR) of crops and weeds. To identify the clades ABC transporters that are associated with NTSR belong, a literature survey was conducted. This survey identified genes that have been associated with NTSR because the mRNA levels of the ABC transporter mRNA are higher in resistant than sensitive populations. These genes are classified as "candidate NTSR genes" and are listed in (Table 2). ABC transporters demonstrated to increase herbicide resistance when expressed at high levels as a transgene or by other experimental methods are classified as "confirmed NTSR genes" and are listed in (Table 3).

We identified a total of 41 ABC transporter candidate NTSR genes in a variety of weed and non-weed species from the literature (Table 2). They include the grass weeds, Italian ryegrass (*Lolium multiflorum*) [65], red sprangletop (*Leptochloa chinensis*) [66], American slough grass (*Beckmannia syzigachne*) [67], shortawn foxtail (*Alopecurus aequalis*) [68] and jungle rice (*Echinochloa colona*) [69], the broad leaf (eudicot) weeds flixweed (*Descurainia Sophia*) [70], hairy fleabane (*Conyza bonariensis*) [71], giant chickweed (*Myosoton aquaticum*) [66, 72], and Palmer amaranth (*Amarnthus palmeri*) [73] and the crop rice (*Oryza sativa*) [74]. The candidate NTSR genes were identified in populations resistance to 5 different HRAC mode of action herbicide classes, including acetolactate synthase inhibitors, acetyl coA carboxylase inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors, glutamine synthetase inhibitors and auxin mimic/cellulose synthesis inhibitors.

There are 12 monophyletic clades of ABC transporters (A, B, C, D, E/F, G, Cp1, Cp2, Cp3, Cp4, Cp5 and CpCm1). Candidate NTSR genes encoding ABC transporters are found in subfamilies ABCA, ABCB, ABCC, ABCF, ABCG and ABCI (Table 2). Thirteen ABCB and 18 ABCC members were identified as candidate NTSR genes. ABCG proteins constitute by far

Table 2. Plant ABC transporter proteins that have been associated with resistance to herbicides are found within 6 classes identified in our analysis.

SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE	HRAC CLASS	REFERENCE
ABCA	LmABCA7 (Italian ryegrass)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[65]
ABCB	CbABCB1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	MaABCB2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
	MaABCB2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	LcABCB2 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	[66]
	ApABCB2 (Palmer amaranth)	Up-regulated in NTSR plants compared to sensitive plants	glufosinate	10	[73]
	OsABCB4 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	BsABCB10 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	[67]
	AaABCB11 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
	OsABCB11 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	MaABCB19 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	CbABCB25 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	CbABCB27 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	MaABCB29 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
ABCC	DsABCC1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
	CbABCC2 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	MaABCC3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
	MaABCC3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	CbABCC4 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	AaABCC8 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
	MaABCC8 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
	CbABCC8 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	LcABCB8 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	[66]
	CbABCC9 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	CbABCC10 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	MaABCC10 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
	MaABCC10 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	OsABCB10 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	CbABCC12 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	CbABCC13 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	CbABCC14 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	LcABCB15 (red sprangletop)	Up-regulated in NTSR plants compared to sensitive plants	cyhalofop-butyl	1	[66]
ABCF	DsABCF5 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
ABCG	CbABCG7 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	MaABCG8(giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	OsABCG11 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	MaABCG14 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[75]
	MaABCG14 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	Tribenuron-methyl	2	[72]
	OsABCG39 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	EcABCG48 (jungle rice)	regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	[69]
ABCI	CbABCI17 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]

https://doi.org/10.1371/journal.pone.0287356.t002

Table 3. Plant ABC transporter proteins that have been empirically shown to confer resistance or tolerance to xenobiotics are found in 3 classes identified in our analysis.

SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE/XENOBIOTIC	HRAC CLASS	REFERENCE
ABCC	AtABCC1 (*A. thaliana*)	Show this gene encodes a protein capable of transporting GS-conjugated metolachlor	metolachlor	15	[76]
ABCC	EcABCC8 (jungle rice)	Causes resistance when overexpressed in transgenic rice	glyphosate	9	[77]
ABCB	AtABCB1 (*A. thaliana*)	Endogenous overexpression gives resistance to multiple herbiciides	dicamba, oryzalin, pendimethalin, MSMA	4, 3, 3, 0	[78]
ABCG	AtABCG19 (*A. thaliana*)	Endogenous overexpression gives resistance	kanamycin	N/A	[79]
ABCG	AtABCG37 (*A. thaliana*)	GOF mutant more resistant, LOF mutant more sensitive	2,4-D	4	[7]
ABCG	AtABCG39 (*A. thaliana*)	LOF mutant is resistant	paraquat	22	[8]

The position of these genes or their *A. thaliana* homologs has been indicated in the phylogeny in Fig 1.

https://doi.org/10.1371/journal.pone.0287356.t003

the largest monophyletic clade we identified in our phylogenetic analysis, representing approximately 2/3 of all the ABC transporters included. However, relatively few ABCG transporters were identified as candidate NTSR genes. From 4 species, we identified 6 distinct ABCG candidate NTSR genes–ABCG7, ABCG8, ABCG11, ABCG14, ABCG39 and ABCG48. A single candidate NTSR gene was found in the literature from each of the groups ABCA, ABCF and ABCI. Taken together, these data suggest candidate NTSR genes encoding ABC transporters belong to clades that were present in the last common ancestor of the Archaeplastida and that the frequency of candidate NTSR genes is not proportional to ABC transporter clade size.

Six ABC transporter proteins have been experimentally confirmed to confer NTSR in plants (Table 3). These 6 genes belong to 3 different clades–ABCB, ABCC and ABCG. These clades comprise sequences from all 10 species of Archaeplastida included in the analysis, and therefore had already diverged in the last common ancestor of the Archaeplastida. It can be concluded that ABC transporters that have been confirmed to contribute to NTSR belong to clades which were already present in the common ancestor of the Archaeplastida. The ABC transporters confirmed to confer resistance to herbicides did so to a variety of HRAC mode of action groups, including very long-chain fatty acid inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors, auxin mimics, microtubule assembly inhibitors, PSI inhibitors, and a herbicide with an unknown mode of action (MSMA). We therefore show that ABC transporter proteins within the same monophyletic clade confer resistance to herbicides with different modes of action.

UGTs groups associated with herbicide resistance evolved after the divergence of Angiosperms from the Archaeplastida lineage

Many UGT genes have also been reported to contribute to NTSR in populations of resistant weeds in the field. We identified a total of 34 candidate NTSR UGT genes in a variety of weed and non-weed species from the literature (Table 4). They include the grass weeds, Italian ryegrass (*Lolium multiflorum*) [65], American slough grass (*Beckmannia syzigachne*) [67, 74, 80], *Brachypodium hybridum* [81], shortawn foxtail (*Alopecurus aequalis*) [68], Tausch's goat grass (*Aegilops tauschii*) [82], and jungle rice (*Echinochloa colona*) [69], the broad leaf (eudicot)

Table 4. Plant UGTs associated with herbicide resistance are found in 11 classes we have identified in our analysis.

GROUP	FAMILY + SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE	HRAC CLASS	REFERENCE
A	79	LmUGT79 (Italian ryegrass)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[65]
	79	OsUGT79 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	Fenoxaprop-P-ethyl	1	[74]
	79B	DsUGT79B6 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
	91A	DsUGT91A1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
B	89A	CbUGT89A (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
C	90A	BsUGT90A1 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	[67]
	71C	DsUGT71C2 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
D	73B	MaUGT73B3 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[75]
	73B	BhUGT73B5 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	[81]
	73B	AaUGT73B5 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
	73C	AtUGT73C (Tausch's goat grass)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[82]
	73C	BsUGT73C1 (American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	[80]
	73C	AaUGT73C5 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
	73D	EcUGT73D1 (jungle rice)	Up-regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	[69]
	73E	CbUGT73E1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
E	71B	DsUGT71B1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
	88A	BhUGT88A1 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	[81]
F	78D	MaUGT78D2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[75]
	78D	MaUGT78D2 (giant chickweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[72]
	78D	DsUGT78D4 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
G	85A	BhUGT85A2 (*Brachypodium hybridum*)	Up-regulated upon herbicide application	pinoxaden	1	[81]
	85A	BsUGT85A2(American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	[80]
	85C	CbUGT85C1 (hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
	85C	CbUGT85C2(hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
H	76B	CbUGT76B1(hairy fleabane)	Up-regulated in NTSR plants compared to sensitive plants	glyphosate	9	[71]
I	83A	AaUGT83A1(shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]

(*Continued*)

Table 4. (Continued)

GROUP	FAMILY + SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE	HRAC CLASS	REFERENCE
	83A	BsUGT83A1(American slough grass)	Up-regulated in NTSR plants compared to sensitive plants	fenoxaprop-p-ethyl	1	[67]
	83A	OsUGT83A1 (rice)	Up-regulated upon safener treatment which causes herbicide tolerance in rice	fenoxaprop-P-ethyl	1	[74]
L	74D	DsUGT74D1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
	75D	AaUGT75D1 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
	75D	EcUGT75D1 (jungle rice)	Up-regulated in NTSR plants compared to sensitive plants	quinclorac	4/29	[69]
	84A	DsUGT84A1 (flixweed)	Up-regulated in NTSR plants compared to sensitive plants	tribenuron-methyl	2	[70]
M	92A	AaUGT92A1 (shortawn foxtail)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[68]
UGT13248		AtUGT13248 (Tausch's goat grass)	Up-regulated in NTSR plants compared to sensitive plants	mesosulfuron-methyl	2	[82]

https://doi.org/10.1371/journal.pone.0287356.t004

weeds flixweed (*Descurainia Sophia*) [70], hairy fleabane (*Conyza bonariensis*) [71], and giant chickweed (*Myosoton aquaticum*) [66, 72] and the crop rice (*Oryza sativa*) [74]. The candidate NTSR genes were identified in populations resistant to 4 classes of HRAC mode of action herbicides, including acetolactate synthase inhibitors, acetyl coA carboxylase inhibitors, enolpyruvyl shikimate phosphate synthase inhibitors and auxin mimic/cellulose synthesis inhibitors.

There are 16 monophyletic groups of UGT sequences in angiosperms. We identified candidate NTSR genes from 11 of these monophyletic groups in angiosperm weeds and crops (Table 4). There are 8 candidate NTSR group D UGT genes in 7 different species. These represent 7 distinct genes, UGT73B3, UGT73B5, UGT73C, UGT73C1, UGT73C5, UGT73D1 and UGT73E1. UGT73B5 was identified in both *B. hybridum* and *A. aequalis*. A single candidate

Table 5. Plant UGTs that have been empirically shown to confer resistance in plants belong to 4 classes identified in our analysis.

GROUP	FAMILY + SUB-FAMILY	GENE NAME	EVIDENCE	HERBICIDE/XENOBIOTIC	HRAC CLASS	REFERENCE
A	91C	AtUGT91C1 (*A. thaliana*)	Demonstrated glucosylation of sulcotrione and herbicide resistance in overexpression lines	sulcotrione	27	[83]
E	72B	AtUGT72B1 (*A. thaliana*)	Knockout plants no longer conjugated these chemicals	3,4-dichloroaniline (DCA) and 2,4,5-trichlorophenol (TCP)	N/A	[84]
D	73C	BsUGT73C1 (American slough grass)	Identified in RNA-seq experiment–causes resistance when expressed in *B. distachyon*	fenoxaprop-p-ethyl	1	[80]
G	85A	BsUGT85A2 (American slough grass)	Identified in RNA-seq experiment–causes resistance when expressed in *B. distachyon*	fenoxaprop-p-ethyl	1	[80]
D	73B	AtUGT73B1 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	[85]
D	73B	AtUGT73B2 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	[85]
D	73B	AtUGT73B3 (*A. thaliana*)	Mutations leading to overexpression cause herbicide resistance	paraquat	22	[85]

The position of these genes or their *A. thaliana* homologs has been indicated in the phylogeny in Fig 3.

https://doi.org/10.1371/journal.pone.0287356.t005

NTSR gene was identified in flixweed, UGT71B1, a group E UGT. Group E is the second largest UGT group. No candidate NTSR UGT sequences were identified from group O, which is the largest UGT group in which angiosperm sequences were identified. These data indicate UGT proteins that detoxify herbicides belong to clades that were present in the last common ancestor of the Angiosperms.

Seven UGT genes, from 4 groups–A, D, E and G–have been experimentally confirmed to confer NTSR in plants (Table 5). There are genes from only two species, the angiosperms *A. thaliana* and *O. sativa*, in each of these groups in our phylogenetic analysis. This suggests that these groups were present in the last common ancestor and evolved following the divergence of the angiosperms from the lycophytes. 6 of the 7 genes identified confer NTSR to 3 HRAC mode of action classes, HPPD inhibitors, acetyl coA inhibitors and PSI inhibitors.

Discussion

Herbicide use has increased continuously since the 1940s, leading to the evolution of herbicide resistance in weeds through selection of naturally occurring alleles that confer herbicide resistance. Overexpression of genes encoding some members of the ATP-binding cassette (ABC) transporter family and uridine diphosphate (UDP) glycosyltransferase (UGT) family in weeds confers herbicide non-target site resistance (NTSR) [2, 4, 5]. ABC transporters are ATP-driven pumps that transport molecules across membranes, while UGT enzymes catalyse the metabolism of a diversity of compounds by transferring a sugar moiety to an acceptor molecule. By carrying out phylogenetic analyses, we set out to resolve the evolutionary relationships of these protein families, and to classify those genes associated with NTSR. We compared sequences from 10 species which represent distinct lineages of the Archaeplastida and generated phylogenetic trees. From the topology of these trees, it can be concluded that ABC transporters were present in the last common ancestor of the Archaeplastida. There are 12 monophyletic clades of ABC transporters (A, B, C, D, E/F, G, Cp1, Cp2, Cp3, Cp4, Cp5, and CpCm1) and one polyphyletic clade (I), all of which existed in the last common ancestor of the Archaeplastida. UGT enzymes existed in the last common ancestor of the streptophytes and have since diversified into 20 monophyletic clades (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Kn1, Pp1, Sm1, and Sm2). Overall, the phylogenetic analysis shows that ABC transporters were present in the last common ancestor of the Archaeplastida, while UGTs proteins are only present in the streptophyte lineage among the Archaeplastida.

In plants, ABC transporters transport of a variety of molecules across membranes [10, 80]. ABC transporters confer herbicide resistance by sequestering herbicides and their metabolites in the vacuole or extracellular spaces, where they cannot interact with their target [16]. We hypothesised that a high diversity of ABC transporters among the Archaeplastida may facilitate the high diversity of substrates which can be recognised and transported by ABC transporters. However, we have shown that closely related ABC transporters, within the same clade, can be associated with resistance to a diverse range of herbicide mode of actions. For example, proteins from clade A are associated with resistance to 6 different herbicides, from 4 different HRAC mode of action groups (HRAC group 1, 2, 9 and 10). This suggests that ABC transporter diversity is not necessary for a diversity of molecules which can be recognised and transported. Nonetheless, this diversity may play a role diversity of ABC transporter localisation, and hence in herbicide detoxification by allowing ABC transporters to sequester chemicals in various locations. These locations include the vacuole and extracellular spaces, where herbicides can be sequestered and are no longer toxic to the plant.

ABC transporter clades in extant plants are ancient and we show that the common ancestor of the Archaeplastida, which existed around 1.6 billion years ago in the Proterozoic era,

contained all clades of ABC transporters identified in our analysis [38, 86, 87]. These ABC transporters evolved during the Proterozoic to carry out various cellular functions, unrelated to herbicide resistance. However, certain ABC transporter genes are now being selected in modern agricultural systems for a new purpose–conferring resistance to herbicides. For example, ABCB1, a P-glycoprotein, which has been shown to cause resistance to dicamba, oryzalin, pendimethalin and MSMA when endogenously overexpressed in *A*. *thaliana* has been characterised [78]. ABCB1 has been shown to be involved in the transport of auxin, a crucial plant hormone essential for numerous developmental processes. ABCB1 mediates auxin efflux, regulating its distribution and concentration in different plant tissues. ABCB1 mutants have altered auxin distribution and as a result, show various developmental defects [88–90]. Herbicide resistance evolving through the altered function of genes that evolved to carry out an entirely different function, can be considered exaptation [91]. Therefore, we argue that the involvement of ABC transporters in herbicide resistance is an example of exaptation.

UGT enzymes catalyse the conjugation of sugars moieties to a variety of acceptor molecules, including herbicides which can lead to detoxification [22]. High sequence diversity among UGTs genes leads to protein structural heterogeneity, which in turn facilitates a wide range of acceptor molecules, including herbicides [24]. We have shown that a large proportion of UGT gene diversity evolved in the angiosperm lineage, and this may account for the ability of UGTs to detoxify herbicides. However, we have also shown that UGTs from closely related clades, or even those from within the same clade, may confer resistance to herbicides with a variety of different chemistries and modes of action. For example, UGT proteins from group D have been associated with, or confer resistance to, 7 different herbicides from 5 different HRAC mode of action groups (HRAC group 1, 2, 4/29, 9 and 22). This suggests sufficient variability has also evolved within UGT clades in extant plants to give UGTs the capacity to conjugate sugars to a diverse range of acceptor molecules, which includes herbicides.

UGTs proteins were present in the last common ancestor of the streptophytes, which existed 1.1–1.2 billion years ago, in the Proterozoic era [92–94]. However, UGTs are not present in the earlier diverging glaucophyte, rhodophyte or chlorophyte lineages. Since UGTs are present in other eukaryotes and some bacteria, their absence in non-streptophyte algal lineages suggests either a loss of UGT genes in these lineages or that UGTs were acquired in the last common ancestor of the streptophyte lineage through horizontal gene transfer. Our data suggest that there was little diversification of UGT sequences until after the vascular plants diverged from other streptophytes in the Silurian period of the Paleozoic era (420–450 million years ago) [92–95]. 17 of the 20 UGT clades we identified, evolved in the vascular lineage. 14 of these groups are present only in angiosperms, suggesting they evolved after the divergence of seed plants from lycophytes, which occurred in the Devonian period of the Paleozoic era (360–400 million years ago) [22, 92]. This diversification of UGTs likely accompanied an increase in morphological complexity of plants and the colonisation of a wider range of terrestrial habitats. UGTs were likely selected to carry out a variety of cellular functions in streptophytes, during their colonisation of land, and subsequent evolution of the vascular plants. Over the last 50 years, they have been selected for herbicide resistance in the agricultural environment.

We have shown that both ABC transporters and UGT enzymes evolved in the Proterozoic era. ABC transporters genes were present in the last common ancestor of the Archaeplastida and had already diversified into many clades in this lineage. Conversely, our data shows that UGTs genes are only present in the streptophyte lineage and diversified mainly in the vascular lineage among plants. The clade diversity of ABC transporters and UGT enzymes in extant plants may contribute to their ability to confer herbicide resistance. We propose that ABC transporters and UGTs were selected to carry out a variety of cellular functions in the Proterozoic, and now some members of these protein families are being selected to carry out an

entirely different function–conferring resistance to herbicides. ABC transporters and UGTs conferring herbicide resistance is therefore an example of exaptation.

Supporting information

S1 Fig. Protein features of ABC transporters and UDP-glycosyltransferases in plants. Diagrams of the typical ABC transporter (A) and UDP-glycosyltransferases (B). A shows the domain architecture of ABC transporters from each of the major sub-families in plants. The diagram is adapted from https://doi.org/10.1038/srep16724. B shows α-helices (green), β-strands (purple) and residues reported to form part of the acceptor pocket (red dashed lines). The PSPG motif reported to interact with the UDP-sugar donor is highlighted in grey. The structure is based on that of AtUGT72B1 reported in https://doi.org/10.1016/j.phytochem.2008.12.009.
(TIF)

S2 Fig. Number of ABC proteins identified in each clade.
(TIF)

S3 Fig. Number of UGT proteins identified in each clade. Number of Family 28 glycosyltransferases identified in each species is also given.
(TIF)

S1 Text. Untrimmed alignment of all ABC transporter sequences used in the phylogenetic analysis.
(TXT)

S2 Text. Untrimmed alignment of all UGT sequences used in the phylogenetic analysis.
(TXT)

S3 Text. Manually trimmed alignment of all ABC transporter sequences used in the phylogenetic analysis.
(TXT)

S4 Text. Manually trimmed alignment of all UGT sequences used in the phylogenetic analysis.
(TXT)

Author Contributions

Writing – original draft: Samuel Caygill.

Writing – review & editing: Liam Dolan.

References

1. Heap I. The International Herbicide-Resistant Weed Database 2023. http://www.weedscience.org/Home.aspx (accessed January 26, 2023).
2. Powles SB, Yu Q. Evolution in Action: Plants Resistant to Herbicides. Annu Rev Plant Biol 2010; 61:317–47. https://doi.org/10.1146/annurev-arplant-042809-112119 PMID: 20192743
3. Yuan JS, Tranel PJ, Stewart CN. Non-target-site herbicide resistance: a family business. Trends Plant Sci 2007; 12:6–13. https://doi.org/10.1016/j.tplants.2006.11.001 PMID: 17161644
4. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A, et al. Mechanisms of evolved herbicide resistance. J Biol Chem 2020; 295:10307–30. https://doi.org/10.1074/jbc.REV120.013572 PMID: 32430396

5. Délye C, Duhoux A, Pernin F, Riggins CW, Tranel PJ. Molecular Mechanisms of Herbicide Resistance. Weed Sci 2015; 63:91–115. https://doi.org/10.1614/WS-D-13-00096.1.
6. Ghanizadeh H, Harrington KC. Non-target Site Mechanisms of Resistance to Herbicides. CRC Crit Rev Plant Sci 2017; 36:24–34. https://doi.org/10.1080/07352689.2017.1316134.
7. Ito H, Gray WM. A gain-of-function mutation in the Arabidopsis pleiotropic drug resistance transporter PDR9 confers resistance to auxinic herbicides. Plant Physiol 2006; 142:63–74. https://doi.org/10.1104/pp.106.084533 PMID: 16877699
8. Xi J, Xu P, Xiang C-B. Loss of AtPDR11, a plasma membrane-localized ABC transporter, confers paraquat tolerance in Arabidopsis thaliana. The Plant Journal 2012; 69:782–91. https://doi.org/10.1111/j.1365-313X.2011.04830.x PMID: 22026747
9. Henikoff S, Greene EA, Pietrokovski S, Bork P, Attwood TK, Hood L. Gene Families: The Taxonomy of Protein Paralogs and Chimeras. Science (1979) 1997; 278:609–14. https://doi.org/10.1126/science.278.5338.609 PMID: 9381171
10. Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, et al. Plant ABC Transporters. The Arabidopsis Book / American Society of Plant Biologists 2011; 9:e0153. https://doi.org/10.1199/tab.0153 PMID: 22303277
11. Theodoulou FL. Plant ABC transporters. Biochimica et Biophysica Acta (BBA)—Biomembranes 2000; 1465:79–103. https://doi.org/10.1016/s0005-2736(00)00132-2 PMID: 10748248
12. Theodoulou FL, Kerr ID. ABC transporter research: going strong 40 years on. Biochem Soc Trans 2015; 43:1033–40. https://doi.org/10.1042/BST20150139 PMID: 26517919
13. Borghi L, Kang J, Ko D, Lee Y, Martinoia E. The role of ABCG-type ABC transporters in phytohormone transport. Biochem Soc Trans 2015; 43:924–30. https://doi.org/10.1042/BST20150106 PMID: 26517905
14. Martinoia E, Klein M, Geisler M, Bovet L, Forestier C, Kolukisaoglu Ü, et al. Multifunctionality of plant ABC transporters–more than just detoxifiers. Planta 2002; 214:345–55. https://doi.org/10.1007/s004250100661 PMID: 11855639
15. Martinoia E, Grill E, Tommasini R, Kreuz K, Amrhein N. ATP-dependent glutathione S-conjugate "export" pump in the vacuolar membrane of plants. Nature 1993; 364:247–9. https://doi.org/10.1038/364247a0.
16. Ge X, d'Avignon DA, Ackerman JJ, Sammons RD. Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. Pest Manag Sci 2010; 66:345–8. https://doi.org/10.1002/ps.1911 PMID: 20063320
17. Higgins CF. ABC TRANSPORTERS: From Microorganisms to Man. vol. 8. 1992.
18. Locher KP. Mechanistic diversity in ATP-binding cassette (ABC) transporters. Nat Struct Mol Biol 2016; 23:487–93. https://doi.org/10.1038/nsmb.3216 PMID: 27273632
19. Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, et al. Plant ABC proteins–a unified nomenclature and updated inventory. Trends Plant Sci 2008; 13:151–9. https://doi.org/10.1016/j.tplants.2008.02.001 PMID: 18299247
20. Saha J, Sengupta A, Gupta K, Gupta B. Molecular phylogenetic study and expression analysis of ATP-binding cassette transporter gene family in Oryza sativa in response to salt stress. Comput Biol Chem 2015; 54:18–32. https://doi.org/10.1016/j.compbiolchem.2014.11.005 PMID: 25531538
21. Yu J, Hu F, Dossa K, Wang Z, Ke T. Genome-wide analysis of UDP-glycosyltransferase super family in Brassica rapa and Brassica oleracea reveals its evolutionary history and functional characterization. BMC Genomics 2017; 18:474. https://doi.org/10.1186/s12864-017-3844-x PMID: 28645261
22. Dimunová D, Matoušková P, Podlipná R, Boušová I, Skálová L. The role of UDP-glycosyltransferases in xenobioticresistance. Drug Metab Rev 2022; 54:282–98. https://doi.org/10.1080/03602532.2022.2083632 PMID: 35635097
23. Yonekura-Sakakibara K, Hanada K. An evolutionary view of functional diversity in family 1 glycosyltransferases. The Plant Journal 2011; 66:182–93. https://doi.org/10.1111/j.1365-313X.2011.04493.x PMID: 21443631
24. Bowles D, Isayenkova J, Lim E-K, Poppenberger B. Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol 2005; 8:254–63. https://doi.org/10.1016/j.pbi.2005.03.007 PMID: 15860422
25. Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S. A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. The Plant Journal 2012; 69:1030–42. https://doi.org/10.1111/j.1365-313X.2011.04853.x PMID: 22077743
26. Ross J, Li Y, Lim E-K, Bowles DJ. Higher plant glycosyltransferases. Genome Biol 2001; 2:reviews3004.1. https://doi.org/10.1186/gb-2001-2-2-reviews3004 PMID: 11182895

27. Gloster TM. Advances in understanding glycosyltransferases from a structural perspective. Curr Opin Struct Biol 2014; 28:131–41. https://doi.org/10.1016/j.sbi.2014.08.012 PMID: 25240227

28. Meech R, Mackenzie PI. STRUCTURE AND FUNCTION OF URIDINE DIPHOSPHATE GLUCURONOSYLTRANSFERASES. Clin Exp Pharmacol Physiol 1997; 24:907–15. https://doi.org/10.1111/j.1440-1681.1997.tb02718.x PMID: 9406655

29. Tiwari P, Sangwan RS, Sangwan NS. Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. Biotechnol Adv 2016; 34:714–39. https://doi.org/10.1016/j.biotechadv.2016.03.006 PMID: 27131396

30. Loutre C, Dixon DP, Brazier M, Slater M, Cole DJ, Edwards R. Isolation of a glucosyltransferase from Arabidopsis thaliana active in the metabolism of the persistent pollutant 3,4-dichloroaniline. The Plant Journal 2003; 34:485–93. https://doi.org/10.1046/j.1365-313x.2003.01742.x PMID: 12753587

31. HACKETT JD, SU YOON H, BUTTERFIELD NJ, SANDERSON MJ, BHATTACHARYA D. Plastid Endosymbiosis: Sources and Timing of the Major Events. Evolution of Primary Producers in the Sea 2007:109–32. https://doi.org/10.1016/B978-012370518-1/50008-4.

32. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A Molecular Timeline for the Origin of Photosynthetic Eukaryotes. Mol Biol Evol 2004; 21:809–18. https://doi.org/10.1093/molbev/msh075 PMID: 14963099

33. Butterfield NJ. Organic preservation of non-mineralizing organisms and the taphonomy of the Burgess Shale. Paleobiology 1990; 16:272–86. https://doi.org/10.1017/S0094837300009994.

34. Nie Y, Foster CSP, Zhu T, Yao R, Duchêne DA, Ho SYW, et al. Accounting for Uncertainty in the Evolutionary Timescale of Green Plants Through Clock-Partitioning and Fossil Calibration Strategies. Syst Biol 2020; 69:1–16. https://doi.org/10.1093/sysbio/syz032 PMID: 31058981

35. Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, et al. The timescale of early land plant evolution. Proceedings of the National Academy of Sciences 2018; 115:E2274–83. https://doi.org/10.1073/pnas.1719588115 PMID: 29463716

36. Becker B. Snow ball earth and the split of Streptophyta and Chlorophyta. Trends Plant Sci 2013; 18:180–3. https://doi.org/10.1016/j.tplants.2012.09.010 PMID: 23102566

37. Herron MD, Hackett JD, Aylward FO, Michod RE. Triassic origin and early radiation of multicellular volvocine algae. Proceedings of the National Academy of Sciences 2009; 106:3254–8. https://doi.org/10.1073/pnas.0811205106 PMID: 19223580

38. Hedges S, Blair JE, Venturi ML, Shoe JL. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol 2004; 4:2. https://doi.org/10.1186/1471-2148-4-2 PMID: 15005799

39. Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? Proceedings of the National Academy of Sciences 2004; 101:15386–91. https://doi.org/10.1073/pnas.0403984101 PMID: 15494441

40. The Arabidopsis Information Resource (TAIR) n.d. https://www.arabidopsis.org/Blast/index.jsp.

41. Rice Genome Annotation Project n.d. http://rice.plantbiology.msu.edu/ (accessed June 24, 2020).

42. Li F-W, Nishiyama T, Waller M, Frangedakis E, Keller J, Li Z, et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants 2020; 6:259–72. https://doi.org/10.1038/s41477-020-0618-2 PMID: 32170292

43. Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, et al. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nat Commun 2014; 5. https://doi.org/10.1038/ncomms4978 PMID: 24865297

44. Phytozome v12.1: Search for genes, families and sequences n.d. https://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST (accessed June 24, 2020).

45. Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, et al. Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature 2004; 428:653–7. https://doi.org/10.1038/nature02398 PMID: 15071595

46. Price DC, Goodenough UW, Roth R, Lee J-H, Kariyawasam T, Mutwil M, et al. Analysis of an improved Cyanophora paradoxa genome assembly. DNA Res 2019; 26:287–99. https://doi.org/10.1093/dnares/dsz009 PMID: 31098614

47. Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res 2021; 49:D344–54. https://doi.org/10.1093/nar/gkaa977 PMID: 33156333

48. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 2019; 20:1160–6. https://doi.org/10.1093/bib/bbx108 PMID: 28968734

49. Okonechnikov K, Golosova O, Fursov M, team the U. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 2012; 28:1166–7. https://doi.org/10.1093/bioinformatics/bts091 PMID: 22368248

50. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010; 59:307–21. https://doi.org/10.1093/sysbio/syq010 PMID: 20525638

51. FigTree n.d. http://tree.bio.ed.ac.uk/software/figtree/ (accessed January 24, 2023).

52. Draw Freely | Inkscape n.d. https://inkscape.org/ (accessed January 24, 2023).

53. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 2000; 408:796–815. https://doi.org/10.1038/35048692 PMID: 11130711

54. Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, Mccombie WR, Ouyang S, et al. Improvement of the oryza sativa nipponbare reference genome using next generation sequence and optical map data. Rice 2013; 6:3–10. https://doi.org/10.1186/1939-8433-6-4.

55. Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, DePamphilis C, et al. The selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science (1979) 2011; 332:960–3. https://doi.org/10.1126/science.1203810 PMID: 21551031

56. Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, et al. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science (1979) 2008; 319:64–9. https://doi.org/10.1126/science.1150646 PMID: 18079367

57. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, et al. Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. Cell 2017; 171:287–304.e15. https://doi.org/10.1016/j.cell.2017.09.030 PMID: 28985561

58. Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, et al. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science (1979) 2007; 318:245–51. https://doi.org/10.1126/science.1143609 PMID: 17932292

59. Zhang J, Fu X-X, Li R-Q, Zhao X, Liu Y, Li M-H, et al. The hornwort genome and early land plant evolution. Nat Plants 2020; 6:107–18. https://doi.org/10.1038/s41477-019-0588-4 PMID: 32042158

60. Hansen SF, Harholt J, Oikawa A, Scheller H V. Plant Glycosyltransferases Beyond CAZy: A Perspective on DUF Families. Front Plant Sci 2012; 3:59. https://doi.org/10.3389/fpls.2012.00059 PMID: 22629278

61. Drula E, Garron M-L, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 2022; 50:D571–7. https://doi.org/10.1093/nar/gkab1045 PMID: 34850161

62. Fan B, Chen T, Zhang S, Wu B, He B. Mining of efficient microbial UDP-glycosyltransferases by motif evolution cross plant kingdom for application in biosynthesis of salidroside n.d. https://doi.org/10.1038/s41598-017-00568-z.

63. Zhang P, Zhang Z, Zhang L, Wang J, Wu C. Glycosyltransferase GT1 family: Phylogenetic distribution, substrates coverage, and representative structural features. Comput Struct Biotechnol J 2020; 18:1383. https://doi.org/10.1016/j.csbj.2020.06.003 PMID: 32637037

64. Foflonker F, Price DC, Qiu H, Palenik B, Wang S, Bhattacharya D. Genome of the halotolerant green alga *P icochlorum* sp. reveals strategies for thriving under fluctuating environmental conditions. Environ Microbiol 2015; 17:412–26. https://doi.org/10.1111/1462-2920.12541.

65. Cechin J, Piasecki C, Benemann DP, Kremer FS, Galli V, Maia LC, et al. Transcriptome Analysis Identifies Candidate Target Genes Involved in Glyphosate-Resistance Mechanism in Lolium multiflorum. Plants (Basel) 2020; 9. https://doi.org/10.3390/plants9060685 PMID: 32481698

66. Chen K, Peng Y, Zhang L, Wang L, Mao D, Zhao Z, et al. Whole transcriptome analysis resulted in the identification of Chinese sprangletop (Leptochloa chinensis) genes involved in cyhalofop-butyl tolerance. BMC Genomics 2021; 22:521. https://doi.org/10.1186/s12864-021-07856-z PMID: 34238252

67. Bai S, Zhao Y, Zhou Y, Wang M, Li Y, Luo X, et al. Identification and expression of main genes involved in non-target site resistance mechanisms to fenoxaprop-p-ethyl in *Beckmannia syzigachne*. Pest Manag Sci 2020; 76:2619–26. https://doi.org/10.1002/ps.5800.

68. Zhao N, Li W, Bai S, Guo W, Yuan G, Wang F, et al. Transcriptome Profiling to Identify Genes Involved in Mesosulfuron-Methyl Resistance in Alopecurus aequalis. Front Plant Sci 2017; 8:1391. https://doi.org/10.3389/fpls.2017.01391 PMID: 28848590

69. Rangani G, Rouse CE, Saski C, Noorai RE, Shankar V, Lawton-Rauh AL, et al. High Resistance to Quinclorac in Multiple-Resistant Echinochloa colona Associated with Elevated Stress Tolerance Gene Expression and Enriched Xenobiotic Detoxification Pathway. Genes (Basel) 2022; 13:515. https://doi.org/10.3390/genes13030515 PMID: 35328069

70. Yang Q, Deng W, Li X, Yu Q, Bai L, Zheng M. Target-site and non-target-site based resistance to the herbicide tribenuron-methyl in flixweed (Descurainia sophia L.). BMC Genomics 2016; 17:551. https://doi.org/10.1186/s12864-016-2915-8 PMID: 27495977

71. Piasecki C, Yang Y, Benemann DP, Kremer FS, Galli V, Millwood RJ, et al. Transcriptomic Analysis Identifies New Non-Target Site Glyphosate-Resistance Genes in Conyza bonariensis. Plants 2019; 8:157. https://doi.org/10.3390/plants8060157 PMID: 31181629

72. Bai S, Liu W, Wang H, Zhao N, Jia S, Zou N, et al. Enhanced Herbicide Metabolism and Metabolic Resistance Genes Identified in Tribenuron-Methyl Resistant *Myosoton aquaticum* L. J Agric Food Chem 2018; 66:9850–7. https://doi.org/10.1021/acs.jafc.8b02740.

73. Salas-Perez RA, Saski CA, Noorai RE, Srivastava SK, Lawton-Rauh AL, Nichols RL, et al. RNA-Seq transcriptome analysis of *Amaranthus palmeri* with differential tolerance to glufosinate herbicide. PLoS One 2018; 13:1–33. https://doi.org/10.1371/journal.pone.0195488.

74. Zhao Y, Li W, Sun L, Wu R, Xu H, Su W, et al. Candidate Genes Involved in Tolerance to Fenoxaprop-P-Ethyl in Rice Induced by Isoxadifen-Ethyl Hydrolysate. Agronomy 2023; 13:225. https://doi.org/10.3390/agronomy13010225.

75. Liu W, Bai S, Zhao N, Jia S, Li W, Zhang L, et al. Non-target site-based resistance to tribenuron-methyl and essential involved genes in Myosoton aquaticum (L.). BMC Plant Biol 2018; 18:225. https://doi.org/10.1186/s12870-018-1451-x PMID: 30305027

76. Lu Y-P, Li Z-S, Rea PA. *AtMRP1* gene of *Arabidopsis* encodes a glutathione *S*-conjugate pump: Isolation and functional definition of a plant ATP-binding cassette transporter gene. Proceedings of the National Academy of Sciences 1997; 94:8243–8. https://doi.org/10.1073/pnas.94.15.8243.

77. Pan L, Yu Q, Wang J, Han H, Mao L, Nyporko A, et al. An ABCC-type transporter endowing glyphosate resistance in plants. Proceedings of the National Academy of Sciences 2021; 118:e2100136118. https://doi.org/10.1073/pnas.2100136118 PMID: 33846264

78. Windsor B, Roux SJ, Lloyd A. Multiherbicide tolerance conferred by AtPgp1 and apyrase overexpression in Arabidopsis thaliana. Nat Biotechnol 2003; 21:428–33. https://doi.org/10.1038/nbt809 PMID: 12640467

79. Mentewab A, biotechnology CS-N, 2005 undefined. Overexpression of an Arabidopsis thaliana ABC transporter confers kanamycin resistance to transgenic plants. NatureCom n.d. https://doi.org/10.1038/nbt1134 PMID: 16116418

80. Pan L, Gao H, Xia W, Zhang T, Dong L. Establishing a herbicide-metabolizing enzyme library in *Beckmannia syzigachne* to identify genes associated with metabolic resistance. J Exp Bot 2016; 67:1745–57. https://doi.org/10.1093/jxb/erv565.

81. Matzrafi M, Shaar-Moshe L, Rubin B, Peleg Z. Unraveling the Transcriptional Basis of Temperature-Dependent Pinoxaden Resistance in Brachypodium hybridum. Front Plant Sci 2017; 8:1064. https://doi.org/10.3389/fpls.2017.01064 PMID: 28680434

82. Zhang D, Li X, Bei F, Jin T, Jia S, Bu R, et al. Investigating the Metabolic Mesosulfuron-Methyl Resistance in *Aegilops tauschii* Coss. By Transcriptome Sequencing Combined with the Reference Genome. J Agric Food Chem 2022; 70:11429–40. https://doi.org/10.1021/acs.jafc.2c04529.

83. Huang X, Zhao S, Zhang Y, Li Y, Shen H, Li X, et al. A novel UDP-glycosyltransferase 91C1 confers specific herbicide resistance through detoxification reaction in Arabidopsis. Plant Physiology and Biochemistry 2021; 159:226–33. https://doi.org/10.1016/j.plaphy.2020.12.026 PMID: 33387851

84. Brazier-Hicks M, Edwards R. Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in Arabidopsis thaliana. The Plant Journal 2005; 42:556–66. https://doi.org/10.1111/j.1365-313X.2005.02398.x PMID: 15860014

85. Lira CE, Ahn JH, Lim J. Molecular Genetic Analysis of Tandemly Located Glycosyltransferase Genes, UGT73B1, UGT73B2, and UGT73B3, in Arabidopsis thaliana. vol. 49. 2006.

86. Archibald JM. Endosymbiosis and Eukaryotic Cell Evolution. Current Biology 2015; 25:R911–21. https://doi.org/10.1016/j.cub.2015.07.055 PMID: 26439354

87. Keeling PJ. The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. Annu Rev Plant Biol 2013; 64:583–607. https://doi.org/10.1146/annurev-arplant-050312-120144 PMID: 23451781

88. Geisler M, Murphy AS. The ABC of auxin transport: the role of p-glycoproteins in plant development. FEBS Lett 2006; 580:1094–102. https://doi.org/10.1016/j.febslet.2005.11.054 PMID: 16359667

89. Blakeslee JJ, Bandyopadhyay A, Lee OR, Mravec J, Titapiwatanakun B, Sauer M, et al. Interactions among PIN-FORMED and P-Glycoprotein Auxin Transporters in *Arabidopsis*. Plant Cell 2007; 19:131–47. https://doi.org/10.1105/tpc.106.040782.

90. Geisler M, Blakeslee JJ, Bouchard R, Lee OR, Vincenzetti V, Bandyopadhyay A, et al. Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1. The Plant Journal 2005; 44:179–94. https://doi.org/10.1111/j.1365-313X.2005.02519.x PMID: 16212599

91. Gould SJ, Vrba ES. Exaptation—a Missing Term in the Science of Form. Paleobiology 1982; 8:4–15. https://doi.org/10.1017/S0094837300004310.

92. Delwiche CF, Cooper ED. The Evolutionary Origin of a Terrestrial Flora. Current Biology 2015; 25: R899–910. https://doi.org/10.1016/j.cub.2015.08.029 PMID: 26439353

93. Becker B, Marin B. Streptophyte algae and the origin of embryophytes. Ann Bot 2009; 103:999–1004. https://doi.org/10.1093/aob/mcp044 PMID: 19273476

94. Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, Delwiche CF, et al. Phylogeny and Molecular Evolution of the Green Algae. CRC Crit Rev Plant Sci 2012; 31:1–46. https://doi.org/10.1080/07352689.2011.615705.

95. Kenrick P, Crane PR. The origin and early evolution of plants on land. Nature 1997; 389:33–9. https://doi.org/10.1038/37918.

References

Abbas, G., Anjum Ali, M., Abbas, Z., Aslam, M., Akram, M., 2009. Impact of different herbicides on broadleaf weeds and yield of wheat, Pak. J. Weed Sci. Res.

Ahsan, N., Lee, D.-G., Lee, K.-W., Alam, I., Lee, S.-H., Bahk, J.D., Lee, B.-H., 2008. Glyphosate-induced oxidative stress in rice leaves revealed by proteomic approach. Plant Physiology and Biochemistry 46, 1062–1070. https://doi.org/10.1016/J.PLAPHY.2008.07.002

Alcántara-de la Cruz, R., Fernández-Moreno, P.T., Ozuna, C. V., Rojano-Delgado, A.M., Cruz-Hipolito, H.E., Domínguez-Valenzuela, J.A., Barro, F., De Prado, R., 2016. Target and Non-target Site Mechanisms Developed by Glyphosate-Resistant Hairy beggarticks (*Bidens pilosa L.)* Populations from Mexico. Front Plant Sci 7, 220549. https://doi.org/10.3389/fpls.2016.01492

Alia, Saradhi, P.P., 1991. Proline Accumulation Under Heavy Metal Stress. J Plant Physiol 138, 554–558. https://doi.org/10.1016/S0176-1617(11)80240-3

The Arabidopsis Genome Initiative., 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815. https://doi.org/10.1038/35048692

Archibald, J.M., 2015. Endosymbiosis and Eukaryotic Cell Evolution. Current Biology 25, R911–R921. https://doi.org/10.1016/J.CUB.2015.07.055

Baerson, S.R., Rodriguez, D.J., Biest, N.A., Tran, M., You, J., Kreuger, R.W., Dill, G.M., Pratley, J.E., Gruys, K.J., 2002. Investigating the mechanism of glyphosate resistance in rigid ryegrass (*Lolium ridigum*). Weed Sci 50, 721–730. https://doi.org/10.1614/0043-1745(2002)050[0721:ITMOGR]2.0.CO;2

Bai, S., Liu, W., Wang, H., Zhao, N., Jia, S., Zou, N., Guo, W., Wang, J., 2018. Enhanced Herbicide Metabolism and Metabolic Resistance Genes Identified in Tribenuron-Methyl Resistant *Myosoton aquaticum* L. J Agric Food Chem 66, 9850–9857. https://doi.org/10.1021/acs.jafc.8b02740

Bai, S., Zhao, Y., Zhou, Y., Wang, M., Li, Y., Luo, X., Li, L., 2020. Identification and expression of main genes involved in non-target site resistance mechanisms to fenoxaprop-p-ethyl in *Beckmannia syzigachne*. Pest Manag Sci 76, 2619–2626. https://doi.org/10.1002/ps.5800

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., DePamphilis, C., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A., Ashton, N.W., Axtell, M.J., Barker, E., Barker, M.S., Bennetzen, J.L., Bonawitz, N.D., Chapple, C., Cheng, C., Correa, L.G.G., Dacre, M., DeBarry, J., Dreyer, I., Elias, M., Engstrom, E.M., Estelle, M., Feng, L., Finet, C., Floyd, S.K., Frommer, W.B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K.G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T.P., Mikami, K., Miyazaki, S., Morinaga, S.I., TakashiMurata, Mueller-Roeber, B., Nelson, D.R., Obara, M., Oguri, Y., Olmstead, R.G., Onodera, N., Petersen, B.L., Pils, B., Prigge, M., Rensing, S.A., Riaño-Pachón, D.M., Roberts, A.W., Sato, Y., Scheller, H.V., Schulz, B., Schulz, C., Shakirov, E. V., Shibagaki, N., Shinohara, N., Shippen, D.E., Sørensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theißen, G., Ulvskov, P., Wakazuki, S., Weng, J.K., Willats, W.W.G.T., Wipf, D., Wolf, P.G., Yang, L., Zimmer, A.D., Zhu, Q., Mitros, T., Hellsten, U., Loqué, D., Otillar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S.,

Rokhsar, D., Grigoriev, I. V., 2011. The selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science (1979) 332, 960–963. https://doi.org/10.1126/science.1203810
Becker, B., 2013. Snow ball earth and the split of Streptophyta and Chlorophyta. Trends Plant Sci 18, 180–183. https://doi.org/10.1016/J.TPLANTS.2012.09.010
Becker, B., Marin, B., 2009. Streptophyte algae and the origin of embryophytes. Ann Bot 103, 999–1004. https://doi.org/10.1093/aob/mcp044
Beffa, R., Menne, H., Köcher, H., 2019. Herbicide Resistance Action Committee (HRAC): Herbicide Classification, Resistance Evolution, Survey, and Resistance Mitigation Activities, in: Modern Crop Protection Compounds. Wiley, pp. 5–32. https://doi.org/10.1002/9783527699261.ch1
Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
Bentley, R., Haslam, E., Sc Referee, D., 1990. The Shikimate Pathway-A Metabolic Tree with Many Branches. Crit Rev Biochem Mol Biol 25, 307–384. https://doi.org/10.3109/10409239009090615
Blakeslee, J.J., Bandyopadhyay, A., Lee, O.R., Mravec, J., Titapiwatanakun, B., Sauer, M., Makam, S.N., Cheng, Y., Bouchard, R., Adamec, J., Geisler, M., Nagashima, A., Sakai, T., Martinoia, E., Friml, J., Peer, W.A., Murphy, A.S., 2007. Interactions among PIN-FORMED and P-Glycoprotein Auxin Transporters in *Arabidopsis*. Plant Cell 19, 131–147. https://doi.org/10.1105/tpc.106.040782
Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G.A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D.H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Orengo, C.A., Pandurangan, A.P., Rivoire, C., Sigrist, C.J.A., Sillitoe, I., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Bateman, A., Finn, R.D., 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res 49, D344–D354. https://doi.org/10.1093/nar/gkaa977
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170
Borger, C., Hashem, A., 2007. Evaluating the double knockdown technique: sequence, application interval, and annual ryegrass growth stage. Australian Journal of Agricultural Research 265–271.
Borghi, L., Kang, J., Ko, D., Lee, Y., Martinoia, E., 2015. The role of ABCG-type ABC transporters in phytohormone transport. Biochem Soc Trans 43, 924–930. https://doi.org/10.1042/BST20150106
Bowles, D., Isayenkova, J., Lim, E.-K., Poppenberger, B., 2005. Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol 8, 254–263. https://doi.org/10.1016/J.PBI.2005.03.007
Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S.S., Althoff, F., Araki, T., Arteaga-Vazquez, M.A., Balasubrmanian, S., Barry, K., Bauer, D., Boehm, C.R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K.M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A.E., Eklund, D.M., Florent, S.N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J.,

Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A.J., Higo, A., Hirakawa, Y., Hundley, H.N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopischke, S., Kubo, M., Kyozuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A.M., Lu, C.-W., De Luna, E., Martienssen, R.A., Minamino, N., Mizutani, Masaharu, Mizutani, Miya, Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M.W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J.M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., Schmutz, J., 2017. Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. Cell 171, 287-304.e15. https://doi.org/10.1016/j.cell.2017.09.030

Brazier, M., Cole, D.J., Edwards, R., 2002. 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. https://doi.org/10.1016/S0031-9422(01)00458-7

Brazier-Hicks, M., Edwards, R., 2005. Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in *Arabidopsis thaliana*. The Plant Journal 42, 556–566. https://doi.org/10.1111/j.1365-313X.2005.02398.x

Burnet, M.W.M., Loveys, B.R., Holtum, J.A.M., Powles, S.B., 1993. Increased Detoxification Is a Mechanism of Simazine Resistance in *Lolium rigidum*. Pestic Biochem Physiol 46, 207–218. https://doi.org/10.1006/PEST.1993.1052

Busi, R., Vila-Aiub, M.M., Powles, S.B., 2011. Genetic control of a cytochrome P450 metabolism-based herbicide resistance mechanism in *Lolium rigidum*. Heredity (Edinb) 106, 817–824. https://doi.org/10.1038/hdy.2010.124

Butterfield, N.J., 1990. Organic preservation of non-mineralizing organisms and the taphonomy of the Burgess Shale. Paleobiology 16, 272–286. https://doi.org/10.1017/S0094837300009994

Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., Martens, S., 2012. A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. The Plant Journal 69, 1030–1042. https://doi.org/10.1111/j.1365-313X.2011.04853.x

Carpenter, E.J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J., Jimenez Vieira, F.R., Bowler, C., Dorrell, R.G., Gitzendanner, M.A., Li, L., Du, W., K. Ullrich, K., Wickett, N.J., Barkmann, T.J., Barker, M.S., Leebens-Mack, J.H., Wong, G.K.-S., 2019. Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). Gigascience 8. https://doi.org/10.1093/gigascience/giz126

Caygill, S., Dolan, L., 2023. ATP binding cassette transporters and uridine diphosphate glycosyltransferases are ancient protein families that evolved roles in herbicide resistance through exaptation. PLoS One 18, e0287356. https://doi.org/10.1371/journal.pone.0287356

Cechin, J., Piasecki, C., Benemann, D.P., Kremer, F.S., Galli, V., Maia, L.C., Agostinetto, D., Vargas, A.L., 2020. Transcriptome Analysis Identifies Candidate Target Genes Involved in Glyphosate-Resistance Mechanism in *Lolium multiflorum*. Plants (Basel) 9. https://doi.org/10.3390/plants9060685

Chen, J., Huang, H., Wei, S., Cui, H., Li, X., Zhang, C., 2020. Glyphosate resistance in *Eleusine indica*: EPSPS overexpression and P106A mutation evolved in the same individuals. Pestic Biochem Physiol 164, 203–208. https://doi.org/10.1016/J.PESTBP.2020.01.014

Chen, K., Peng, Y., Zhang, L., Wang, Long, Mao, D., Zhao, Z., Bai, L., Wang, Lifeng, 2021. Whole transcriptome analysis resulted in the identification of Chinese sprangletop (*Leptochloa chinensis*) genes involved in cyhalofop-butyl tolerance. BMC Genomics 22, 521. https://doi.org/10.1186/s12864-021-07856-z

Christenhusz, M.J.M., Byng, J.W., 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261, 201. https://doi.org/10.11646/phytotaxa.261.3.1

Claussen, W., 2005. Proline as a measure of stress in tomato plants. Plant Science 168, 241–248. https://doi.org/10.1016/J.PLANTSCI.2004.07.039

Cole, D.J., John Owen, W., 1987. Influence of monooxygenase inhibitors on the metabolism of the herbicides chlortoluron and metolachlor in cell suspension cultures. Plant Science 50, 13–20. https://doi.org/10.1016/0168-9452(87)90025-2

Cole, S.P.C., Deeley, R.G., 2006. Transport of glutathione and glutathione conjugates by MRP1. Trends Pharmacol Sci 27, 438–46. https://doi.org/10.1016/j.tips.2006.06.008

Conte, S.S., Lloyd, A.M., 2011. Exploring multiple drug and herbicide resistance in plants—Spotlight on transporter proteins. Plant Science 180, 196–203. https://doi.org/10.1016/J.PLANTSCI.2010.10.015

Cummins, I., Cole, D.J., Edwards, R., 1999. A role for glutathione transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. The Plant Journal 18, 285–292. https://doi.org/10.1046/j.1365-313X.1999.00452.x

Cummins, I., Dixon, D.P., Freitag-Pohl, S., Skipsey, M., Edwards, R., 2011. Multiple roles for plant glutathione transferases in xenobiotic detoxification. Drug Metab Rev 43, 266–280. https://doi.org/10.3109/03602532.2011.552910

Cummins, I., Moss, S., Cole, D.J., Edwards, R., 1997. Glutathione transferases in herbicide-resistant and herbicide-susceptible black-grass (*Alopecurus myosuroides*). Pestic Sci 51, 244–250. https://doi.org/10.1002/(SICI)1096-9063(199711)51:3<244::AID-PS643>3.0.CO;2-2

Cummins, I., Wortley, D.J., Sabbadin, F., He, Z., Coxon, C.R., Straker, H.E., Sellars, J.D., Knight, K., Edwards, L., Hughes, D., Kaundun, S.S., Hutchings, S.-J., Steel, P.G., Edwards, R., 2013. Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. Proc Natl Acad Sci U S A 110, 5812–7. https://doi.org/10.1073/pnas.1221179110

Dan Hess, F., 2018. Herbicide Absorption and Translocation and their Relationship to Plant Tolerances and Susceptibility, in: Weed Physiology. CRC Press, pp. 191–214. https://doi.org/10.1201/9781351077736-8

Danielson, B.S.P., 2002. The Cytochrome P450 Superfamily: Biochemistry, Evolution and Drug Metabolism in Humans. Curr Drug Metab 3, 561–597. https://doi.org/10.2174/1389200023337054

Dayan, F.E., 2019. Current Status and Future Prospects in Herbicide Discovery. Plants (Basel) 8. https://doi.org/10.3390/plants8090341

de Vries, J., Archibald, J.M., 2018. Plant evolution: landmarks on the path to terrestrial life. New Phytologist 217, 1428–1434. https://doi.org/10.1111/nph.14975

Delwiche, C.F., Cooper, E.D., 2015. The Evolutionary Origin of a Terrestrial Flora. Current Biology 25, R899–R910. https://doi.org/10.1016/J.CUB.2015.08.029

Délye, C., Duhoux, A., Pernin, F., Riggins, C.W., Tranel, P.J., 2015. Molecular Mechanisms of Herbicide Resistance. Weed Sci 63, 91–115. https://doi.org/10.1614/WS-D-13-00096.1
Délye, C., Jasieniuk, M., Le Corre, V., 2013. Deciphering the evolution of herbicide resistance in weeds. Trends in Genetics 29, 649–658. https://doi.org/10.1016/J.TIG.2013.06.001
Devine, M.D., Shukla, A., 2000. Altered target sites as a mechanism of herbicide resistance. Crop Protection 19, 881–889. https://doi.org/10.1016/S0261-2194(00)00123-X
Dimaano, N.G., Iwakami, S., 2021. Cytochrome P450 -mediated herbicide metabolism in plants: current understanding and prospects. Pest Manag Sci 77, 22–32. https://doi.org/10.1002/ps.6040
Dimunová, D., Matoušková, P., Podlipná, R., Boušová, I., Skálová, L., 2022. The role of UDP-glycosyltransferases in xenobioticresistance. Drug Metab Rev 54, 282–298. https://doi.org/10.1080/03602532.2022.2083632
Donoghue, P.C.J., Harrison, C.J., Paps, J., Schneider, H., 2021. The evolutionary emergence of land plants. https://doi.org/10.1016/j.cub.2021.07.038
Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F., Philippe, H., 2004. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? Proceedings of the National Academy of Sciences 101, 15386–15391. https://doi.org/10.1073/pnas.0403984101
Dowson, Amanda J, Lloyd, A.J., Cuming, A.C., Roper, D.I., Frigerio, L., Dowson, C.G., 2022. Plant peptidoglycan precursor biosynthesis: Conservation between moss chloroplasts and Gram negative bacteria. Plant Physiol. https://doi.org/10.1093/plphys/kiac176
Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., Terrapon, N., 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 50, D571–D577. https://doi.org/10.1093/nar/gkab1045
Duke, S.O., 2018. The history and current status of glyphosate. Pest Manag Sci 74, 1027–1034. https://doi.org/10.1002/ps.4652
Duke, S.O., 2012. Why have no new herbicide modes of action appeared in recent years? Pest Manag Sci 68, 505–512. https://doi.org/10.1002/ps.2333
Duke, S.O., 2005. Taking stock of herbicide-resistant crops ten years after introduction. Pest Manag Sci 61, 211–218. https://doi.org/10.1002/ps.1024
Duke, S.O., 1988. Glyphosate. Kearney, P.C. and Kaufman, D.D., Eds., Herbicides: Chemistry, Degradation, and Mode of Action, Dekker, New York, 1–70. https://doi.org/10.4236/as.2013.44028
Duke, S.O., Powles, S.B., 2008. Glyphosate: a once-in-a-century herbicide. Pest Manag Sci 64, 319–325. https://doi.org/10.1002/ps.1518
Eceiza, M.V., Gil-Monreal, M., Barco-Antoñanzas, M., Zabalza, A., Royuela, M., 2022. The moderate oxidative stress induced by glyphosate is not detected in *Amaranthus palmeri* plants overexpressing EPSPS. J Plant Physiol 274, 153720. https://doi.org/10.1016/J.JPLPH.2022.153720
El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., Finn, R.D., 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47, D427–D432. https://doi.org/10.1093/nar/gky995
Eschenburg, S., Priestman, M., Schönbrunn, E., 2005. Evidence that the fosfomycin target Cys115 in UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) is essential for product release. J Biol Chem 280, 3757–63. https://doi.org/10.1074/jbc.M411325200

Fan, B., Chen, T., Zhang, S., Wu, B., He, B., n.d. Mining of efficient microbial UDP-glycosyltransferases by motif evolution cross plant kingdom for application in biosynthesis of salidroside. https://doi.org/10.1038/s41598-017-00568-z

Feng, P.C.C., Tran, M., Chiu, T., Sammons, R.D., Heck, G.R., CaJacob, C.A., 2004. Investigations into glyphosate-resistant horseweed (*Conyza canadensis*): retention, uptake, translocation, and metabolism. Weed Sci 52, 498–505. https://doi.org/10.1614/WS-03-137R

Florian Krekel, Claudia Oecking, Nikolaus Amrhein, and, Peter Macheroux., 1999. Substrate and Inhibitor-Induced Conformational Changes in the Structurally Related Enzymes UDP-N-Acetylglucosamine Enolpyruvyl Transferase (MurA) and 5-Enolpyruvylshikimate 3-Phosphate Synthase (EPSPS). https://doi.org/10.1021/BI990412O

Foflonker, F., Price, D.C., Qiu, H., Palenik, B., Wang, S., Bhattacharya, D., 2015. Genome of the halotolerant green alga *Picochlorum* sp. reveals strategies for thriving under fluctuating environmental conditions. Environ Microbiol 17, 412–426. https://doi.org/10.1111/1462-2920.12541

Fonné-Pfister, R., Gaudin, J., Kreuz, K., Ramsteiner, K., Ebert, E., 1990. Hydroxylation of primisulfuron by an inducible cytochrome P450-dependent monooxygenase system from maize. Pestic Biochem Physiol 37, 165–173. https://doi.org/10.1016/0048-3575(90)90122-I

Franco-Ortega, S., Goldberg-Cavalleri, A., Walker, A., Brazier-Hicks, M., Onkokesung, N., Edwards, R., 2021. Non-target Site Herbicide Resistance Is Conferred by Two Distinct Mechanisms in Black-Grass (*Alopecurus myosuroides*). Front Plant Sci 12, 194. https://doi.org/10.3389/fpls.2021.636652

Frear, D.S., 1995. Wheat Microsomal Cytochrome P450 Monooxygenases: Characterization and Importance in the Metabolic Detoxification and Selectivity of Wheat Herbicides. Drug Metabol Drug Interact 12, 329–358. https://doi.org/10.1515/DMDI.1995.12.3-4.329

Frear, D.S., Swanson, H.R., Tanaka, F.S., 1969. N-demethylation of substituted 3-(phenyl)-1-methylureas: Isolation and characterization of a microsomal mixed function oxidase from cotton. Phytochemistry 8, 2157–2169. https://doi.org/10.1016/S0031-9422(00)88175-3

Friedrich, T., 2005. Does No-Till Farming Require More Herbicides? Outlooks on Pest Management 16, 188–191. https://doi.org/DOI:10.1564/16aug12

Funke, T., Han, H., Healy-Fried, M.L., Fischer, M., Schö, E., 2006. Molecular basis for the herbicide resistance of Roundup Ready crops.

Funke, T., Yang, Y., Han, H., Healy-Fried, M., Olesen, S., Becker, A., Schönbrunn, E., 2009. Structural Basis of Glyphosate Resistance Resulting from the Double Mutation Thr97 → Ile and Pro101 → Ser in 5-Enolpyruvylshikimate-3-phosphate Synthase from *Escherichia coli*. Journal of Biological Chemistry 284, 9854–9860. https://doi.org/10.1074/JBC.M809771200

Gaines, T.A., Duke, S.O., Morran, S., G Rigon, C.A., Tranel, P.J., Küpper, A., Dayan, F.E., 2020. Mechanisms of evolved herbicide resistance Downloaded from 7, 2020. https://doi.org/10.1074/jbc.REV120.013572

Gaines, T.A., Lorentz, L., Figge, A., Herrmann, J., Maiwald, F., Ott, M., Han, H., Busi, R., Yu, Q., Powles, S.B., Beffa, R., 2014. RNA -Seq transcriptome analysis to identify genes involved in metabolism-based diclofop resistance in *Lolium rigidum*. The Plant Journal 78, 865–876. https://doi.org/10.1111/tpj.12514

Gaines, T.A., Patterson, E.L., Neve, P., 2019. Molecular mechanisms of adaptive evolution revealed by global selection for glyphosate resistance. New Phytologist 223, 1770–1775. https://doi.org/10.1111/nph.15858

Gaines, T.A., Shaner, D.L., Ward, S.M., Leach, J.E., Preston, C., Westra, P., 2011. Mechanism of Resistance of Evolved Glyphosate-Resistant Palmer Amaranth (*Amaranthus palmeri*). J Agric Food Chem 59, 5886–5889. https://doi.org/10.1021/jf104719k

Gaines, T.A., Zhang, W., Wang, D., Bukun, B., Chisholm, S.T., Shaner, D.L., Nissen, S.J., Patzoldt, W.L., Tranel, P.J., Culpepper, A.S., Grey, T.L., Webster, T.M., Vencill, W.K., Sammons, R.D., Jiang, J., Preston, C., Leach, J.E., Westra, P., 2010. Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. Proceedings of the National Academy of Sciences 107, 1029–1034. https://doi.org/10.1073/pnas.0906649107

García, M.J., Palma-Bautista, C., Rojano-Delgado, A.M., Bracamonte, E., Portugal, J., Alcántara-de la Cruz, R., De Prado, R., 2019. The Triple Amino Acid Substitution TAP-IVS in the EPSPS Gene Confers High Glyphosate Resistance to the Superweed *Amaranthus hybridus*. Int J Mol Sci 20, 2396. https://doi.org/10.3390/ijms20102396

Ge, X., d'Avignon, D.A., Ackerman, J.J., Sammons, R.D., 2010. Rapid vacuolar sequestration: the horseweed glyphosate resistance mechanism. Pest Manag Sci 66, 345–348. https://doi.org/10.1002/ps.1911

Geisler, M., Blakeslee, J.J., Bouchard, R., Lee, O.R., Vincenzetti, V., Bandyopadhyay, A., Titapiwatanakun, B., Peer, W.A., Bailly, A., Richards, E.L., Ejendal, K.F.K., Smith, A.P., Baroux, C., Grossniklaus, U., Müller, A., Hrycyna, C.A., Dudler, R., Murphy, A.S., Martinoia, E., 2005. Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1. The Plant Journal 44, 179–194. https://doi.org/10.1111/j.1365-313X.2005.02519.x

Geisler, M., Murphy, A.S., 2006. The ABC of auxin transport: the role of p-glycoproteins in plant development. FEBS Lett 580, 1094–102. https://doi.org/10.1016/j.febslet.2005.11.054

Gerwick, B.C., 2010. Thirty years of herbicide discovery: surveying the past and contemplating the future, in: Agrow Report. Informa; London, UK, pp. VII–IX.

Ghanizadeh, H., Buddenhagen, C.E., Harrington, K.C., James, T.K., 2019. The Genetic Inheritance of Herbicide Resistance in Weeds. CRC Crit Rev Plant Sci 38, 295–312. https://doi.org/10.1080/07352689.2019.1665769

Ghanizadeh, H., Harrington, K.C., 2017. Non-target Site Mechanisms of Resistance to Herbicides. CRC Crit Rev Plant Sci 36, 24–34. https://doi.org/10.1080/07352689.2017.1316134

Gherekhloo, J., Fernández-Moreno, P.T., Alcántara-de la Cruz, R., Sánchez-González, E., Cruz-Hipolito, H.E., Domínguez-Valenzuela, J.A., De Prado, R., 2017. Pro-106-Ser mutation and EPSPS overexpression acting together simultaneously in glyphosate-resistant goosegrass (*Eleusine indica*). Sci Rep 7, 6702. https://doi.org/10.1038/s41598-017-06772-1

Ghitti, E., Rolli, E., Crotti, E., Borin, S., 2022. Flavonoids Are Intra- and Inter-Kingdom Modulator Signals. Microorganisms 10, 2479. https://doi.org/10.3390/microorganisms10122479

Gianessi, L.P., 2013. The increasing importance of herbicides in worldwide crop production. Pest Manag Sci 69, 1099–1105. https://doi.org/10.1002/ps.3598

Gianessi, L.P., 2008. Economic impacts of glyphosate-resistant crops. Pest Manag Sci 64, 346–352. https://doi.org/10.1002/ps.1490

Gloster, T.M., 2014. Advances in understanding glycosyltransferases from a structural perspective. Curr Opin Struct Biol 28, 131–141. https://doi.org/10.1016/J.SBI.2014.08.012
Gould, S.J., Vrba, E.S., 1982. Exaptation—a Missing Term in the Science of Form. Paleobiology 8, 4–15. https://doi.org/10.1017/S0094837300004310
Green, J.M., 2009. Evolution of Glyphosate-Resistant Crop Technology. Weed Sci 57, 108–117. https://doi.org/10.1614/WS-08-030.1
Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 59, 307–321. https://doi.org/10.1093/sysbio/syq010
Hackett, J.D., Su Yoon, H., Butterfield, N.J., Sanderson, M.J., Bhattacharya, D., 2007. Plastid Endosymbiosis: Sources and Timing of the Major Events. Evolution of Primary Producers in the Sea 109–132. https://doi.org/10.1016/B978-012370518-1/50008-4
Han, H., Vila-Aiub, M.M., Jalaludin, A., Yu, Q., Powles, S.B., 2017. A double EPSPS gene mutation endowing glyphosate resistance shows a remarkably high resistance cost. Plant Cell Environ 40, 3031–3042. https://doi.org/10.1111/pce.13067
Hansen, S.F., Harholt, J., Oikawa, A., Scheller, H. V., 2012. Plant Glycosyltransferases Beyond CAZy: A Perspective on DUF Families. Front Plant Sci 3, 59. https://doi.org/10.3389/fpls.2012.00059
Harker, K.N., O'Donovan, J.T., 2013. Recent Weed Control, Weed Management, and Integrated Weed Management. Weed Technology 27, 1–11. https://doi.org/10.1614/WT-D-12-00109.1
Harper, J.L., 1956. The evolution of weeds in relation to resistance to herbicides. Proceedings of the Third British Weed Control Conference.
Harris, B.J., Clark, J.W., Schrempf, D., Szöllősi, G.J., Donoghue, P.C.J., Hetherington, A.M., Williams, T.A., 2022. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. Nat Ecol Evol 6, 1634–1643. https://doi.org/10.1038/s41559-022-01885-x
Harris, B.J., Harrison, C.J., Hetherington, A.M., Williams, T.A., 2020. Phylogenomic Evidence for the Monophyly of Bryophytes and the Reductive Evolution of Stomata. Current Biology 0. https://doi.org/10.1016/J.CUB.2020.03.048
Hatzios, K.K. ed., 1997. Regulation of enzymatic systems detoxifying xenobiotics in plants (Vol. 37). Springer Science & Business Media. Hawkins, N.J., Bass, C., Dixon, A., Neve, P., 2019. The evolutionary origins of pesticide resistance. Biological Reviews 94, 135–155. https://doi.org/10.1111/brv.12440
Haywood, J., Vadlamani, G., Stubbs, K.A., Mylne, J.S., 2021. Antibiotic resistance lessons for the herbicide resistance crisis. Pest Manag Sci 77, 3807–3814. https://doi.org/10.1002/ps.6357
Healy-Fried, M.L., Funke, T., Priestman, M.A., Han, H., Schönbrunn, E., 2007. Structural Basis of Glyphosate Tolerance Resulting from Mutations of Pro101 in *Escherichia coli* 5-Enolpyruvylshikimate-3-phosphate Synthase. Journal of Biological Chemistry 282, 32949–32955. https://doi.org/10.1074/JBC.M705624200
Heap, 2023. The International Herbicide-Resistant Weed Database. [WWW Document]. URL http://www.weedscience.org/home.aspx
Heap, 2014. Global perspective of herbicide-resistant weeds. Pest Manag Sci 70, 1306–1315. https://doi.org/10.1002/ps.3696

Heap, Duke, S.O., 2018. Overview of glyphosate-resistant weeds worldwide. Pest Manag Sci 74, 1040–1049. https://doi.org/10.1002/ps.4760
Heap, 2023. The International Herbicide-Resistant Weed Database [WWW Document].
Hedges, S., Blair, J.E., Venturi, M.L., Shoe, J.L., 2004. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol 4, 2. https://doi.org/10.1186/1471-2148-4-2
Heijenoort, J. v., 2001. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. Glycobiology 11, 25R-36R. https://doi.org/10.1093/glycob/11.3.25R
Henikoff, S., Greene, E.A., Pietrokovski, S., Bork, P., Attwood, T.K., Hood, L., 1997. Gene Families: The Taxonomy of Protein Paralogs and Chimeras. Science (1979) 278, 609–614. https://doi.org/10.1126/science.278.5338.609
Herrmann, K.M., 1995. The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. Plant Cell 7, 907–919. https://doi.org/10.1105/tpc.7.7.907
Herrmann, K.M., Weaver, L.M., 1999. The Shikimate Pathway. Annu Rev Plant Physiol Plant Mol Biol 50, 473–503. https://doi.org/10.1146/annurev.arplant.50.1.473
Herron, M.D., Hackett, J.D., Aylward, F.O., Michod, R.E., 2009. Triassic origin and early radiation of multicellular volvocine algae. Proceedings of the National Academy of Sciences 106, 3254–3258. https://doi.org/10.1073/pnas.0811205106
Higgins, C.F., 1992. ABC Transporters: From Microorganisms to Man, Annu. Rev. Cell Bioi.
Hilton, H., 1957. Herbicide tolerant strains of weeds. Sugar Planters Association Annual Report 69–72.
Hirano, T., Tanidokoro, K., Shimizu, Y., Kawarabayasi, Y., Ohshima, T., Sato, M., Tadano, S., Ishikawa, H., Takio, S., Takechi, K., Takano, H., 2016. Moss Chloroplasts Are Surrounded by a Peptidoglycan Wall Containing D-Amino Acids. Plant Cell 28, 1521–1532. https://doi.org/10.1105/tpc.16.00104
Hofmann, N.R., 2016. Invisible No Longer: Peptidoglycan in Moss Chloroplasts. Plant Cell 28, tpc.00521.2016. https://doi.org/10.1105/tpc.16.00521
Homi, S., Takechi, K., Tanidokoro, K., Sato, H., Takio, S., Takano, H., 2009. The Peptidoglycan Biosynthesis Genes MurA and MraY are Related to Chloroplast Division in the Moss *Physcomitrella patens*. Plant Cell Physiol 50, 2047–2056. https://doi.org/10.1093/pcp/pcp158
Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N., Moriyama, T., Ikeuchi, M., Watanabe, M., Wada, H., Kobayashi, K., Saito, M., Masuda, T., Sasaki-Sekimoto, Y., Mashiguchi, K., Awai, K., Shimojima, M., Masuda, S., Iwai, M., Nobusawa, T., Narise, T., Kondo, S., Saito, H., Sato, R., Murakawa, M., Ihara, Y., Oshima-Yamada, Y., Ohtaka, K., Satoh, M., Sonobe, K., Ishii, M., Ohtani, R., Kanamori-Sato, M., Honoki, R., Miyazaki, D., Mochizuki, H., Umetsu, J., Higashi, K., Shibata, D., Kamiya, Y., Sato, N., Nakamura, Y., Tabata, S., Ida, S., Kurokawa, K., Ohta, H., 2014. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. Nat Commun 5. https://doi.org/10.1038/ncomms4978
Hoss, N.E., Al-Khatib, K., Peterson, D.E., Loughin, T.M., 2003. Efficacy of glyphosate, glufosinate, and imazethapyr on selected weed species. Weed Sci 51, 110–117. https://doi.org/10.1614/0043-1745(2003)051[0110:EOGGAI]2.0.CO;2
Hossain, M.M., 2015. Recent perspective of herbicide: Review of demand and adoption in world agriculture. Journal of the Bangladesh Agricultural University 13, 13–24. https://doi.org/10.22004/ag.econ.211172

HRAC, 2022. HRAC Mode of Action Classification 2022 [WWW Document]. URL https://hracglobal.com/files/HRAC_MOA_Poster_January_6_2022.pdf (accessed 8.29.23).

https://www.verifiedmarketresearch.com/product/glyphosate-market [WWW Document], 2023. URL https://www.verifiedmarketresearch.com/product/glyphosate-market/#:~:text=Glyphosate%20Market%20size%20was%20valued,of%20glyphosate%20throughout%20the%20world. (accessed 8.25.23).

Huang, X., Zhao, S., Zhang, Y., Li, Y., Shen, H., Li, X., Hou, B., 2021. A novel UDP-glycosyltransferase 91C1 confers specific herbicide resistance through detoxification reaction in Arabidopsis. Plant Physiology and Biochemistry 159, 226–233. https://doi.org/10.1016/J.PLAPHY.2020.12.026

Hull, R.I., Moss, S.R., 2012. Is the increasing reliance on residual herbicides for *Alopecurus myosuroides* (black-grass) control sustainable? Asp Appl Biol 25–32.

Ishizaki, K., Chiyoda, S., Yamato, K.T., Kohchi, T., 2008. Agrobacterium-Mediated Transformation of the Haploid Liverwort *Marchantia polymorpha L*., an Emerging Model for Plant Biology. Plant Cell Physiol 49, 1084–1091. https://doi.org/10.1093/pcp/pcn085

Ito, H., Gray, W.M., 2006. A gain-of-function mutation in the Arabidopsis pleiotropic drug resistance transporter PDR9 confers resistance to auxinic herbicides. Plant Physiol 142, 63–74. https://doi.org/10.1104/pp.106.084533

Iwakami, S., Uchino, A., Kataoka, Y., Shibaike, H., Watanabe, H., Inamura, T., 2014. Cytochrome P450 genes induced by bispyribac-sodium treatment in a multiple-herbicide-resistant biotype of *Echinochloa phyllopogon*. Pest Manag Sci 70, 549–558. https://doi.org/10.1002/ps.3572

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., Soltis, D.E., Clifton, S.W., Schlarbaum, S.E., Schuster, S.C., Ma, H., Leebens-Mack, J., dePamphilis, C.W., 2011. Ancestral polyploidy in seed plants and angiosperms. Nature 473, 97–100. https://doi.org/10.1038/nature09916

Jones, P., Vogt, T., 2001. Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. Planta 213, 164–174. https://doi.org/10.1007/s004250000492

Jugulam, M., Shyam, C., 2019. Non-Target-Site Resistance to Herbicides: Recent Developments. Plants 8, 417. https://doi.org/10.3390/plants8100417

Julian O.D. Coleman., 1997. reviews Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. Trends in Plant Sciences 2.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kang, J., Park, J., Choi, H., Burla, B., Kretzschmar, T., Lee, Y., Martinoia, E., 2011. Plant ABC Transporters. Arabidopsis Book 9, e0153. https://doi.org/10.1199/tab.0153

Karavangeli, M., Labrou, N.E., Clonis, Y.D., Tsaftaris, A., 2005. Development of transgenic tobacco plants overexpressing maize glutathione S-transferase I for chloroacetanilide

herbicides phytoremediation. Biomol Eng 22, 121–128. https://doi.org/10.1016/J.BIOENG.2005.03.001
Katayama, N., Takano, H., Sugiyama, M., Takio, S., Sakai, A., Tanaka, K., Kuroiwa, H., Ono, K., 2003. Effects of Antibiotics that Inhibit the Bacterial Peptidoglycan Synthesis Pathway on Moss Chloroplast Division. Plant Cell Physiol 44, 776–781. https://doi.org/10.1093/pcp/pcg096
Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20, 1160–1166. https://doi.org/10.1093/bib/bbx108
Katoh, K., Toh, H., 2010. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics 26, 1899–1900. https://doi.org/10.1093/bioinformatics/btq224
Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., Mccombie, W.R., Ouyang, S., Schwartz, D.C., Tanaka, T., Wu, J., Zhou, S., Childs, K.L., Davidson, R.M., Lin, H., Quesada-Ocampo, L., Vaillancourt, B., Sakai, H., Lee, S.S., Kim, J., Numa, H., Itoh, T., Buell, C.R., Matsumoto, T., 2013. Improvement of the *Oryza sativa* nipponbare reference genome using next generation sequence and optical map data. Rice 6, 3–10. https://doi.org/10.1186/1939-8433-6-4
Keeling, P.J., 2013. The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. Annu Rev Plant Biol 64, 583–607. https://doi.org/10.1146/annurev-arplant-050312-120144
Kenrick, P., Crane, P.R., 1997. The origin and early evolution of plants on land. Nature 389, 33–39. https://doi.org/10.1038/37918
Kenrick, P., Strullu-Derrien, C., 2014. The Origin and Early Evolution of Roots. Plant Physiol 166, 570–580. https://doi.org/10.1104/pp.114.244517
Kishore, G.M. and Shah, D.M., 1988. Amino acid biosynthesis inhibitors as herbicides. Annual review of biochemistry, 57(1), pp.627-663.
Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211–3217. https://doi.org/10.1093/bioinformatics/bts611
Kreiner, J.M., Giacomini, D.A., Bemm, F., Waithaka, B., Regalado, J., Lanz, C., Hildebrandt, J., Sikkema, P.H., Tranel, P.J., Weigel, D., Stinchcombe, J.R., Wright, S.I., 2019. Multiple modes of convergent adaptation in the spread of glyphosate-resistant *Amaranthus tuberculatus*. Proc Natl Acad Sci U S A 116, 21076–21084. https://doi.org/10.1073/pnas.1900870116
Kreiner, J.M., Tranel, P.J., Weigel, D., Stinchcombe, J.R., Wright, S.I., 2021. The genetic architecture and population genomic signatures of glyphosate resistance in *Amaranthus tuberculatus*. Mol Ecol mec.15920. https://doi.org/10.1111/mec.15920
Kreuz, K., Fonné-Pfister, R., 1992. Herbicide-insecticide interaction in maize: Malathion inhibits cytochrome P450-dependent primisulfuron metabolism. Pestic Biochem Physiol 43, 232–240. https://doi.org/10.1016/0048-3575(92)90036-Y
Krishna N. Reddy, Agnes M. Rimando, and, Duke, S.O., 2004. Aminomethylphosphonic Acid, a Metabolite of Glyphosate, Causes Injury in Glyphosate-Treated, Glyphosate-Resistant Soybean. https://doi.org/10.1021/JF049605V
Laforest, M., Soufiane, B., Simard, M.-J., Obeid, K., Page, E., Nurse, R.E., 2017. Acetyl-CoA carboxylase overexpression in herbicide-resistant large crabgrass (*Digitaria sanguinalis*). Pest Manag Sci 73, 2227–2235. https://doi.org/10.1002/ps.4675

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., Forner, J., 2013. GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis. PLoS One 8, e83043. https://doi.org/10.1371/journal.pone.0083043

Leliaert, F., Smith, D.R., Moreau, H., Herron, M.D., Verbruggen, H., Delwiche, C.F., De Clerck, O., 2012. Phylogeny and Molecular Evolution of the Green Algae. CRC Crit Rev Plant Sci 31, 1–46. https://doi.org/10.1080/07352689.2011.615705

Leslie, T., Baucom, R.S., 2014. *De Novo* Assembly and Annotation of the Transcriptome of the Agricultural Weed *Ipomoea purpurea* Uncovers Gene Expression Changes Associated with Herbicide Resistance. G3 Genes|Genomes|Genetics 4, 2035–2047. https://doi.org/10.1534/g3.114.013508

Letouzé, A., Gasquez, J., 2003. Enhanced activity of several herbicide-degrading enzymes: a suggested mechanism responsible for multiple resistance in blackgrass (*Alopecurus myosuroides Huds.*). Agronomie 23, 601–608. https://doi.org/10.1051/agro:2003036

Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49, W293–W296. https://doi.org/10.1093/nar/gkab301

Levin, J.G., Sprinson, D.B., 1964. The Enzymatic Formation and Isolation of 3-Enolpyruvylshikimate 5-Phosphate. Journal of Biological Chemistry 239, 1142–1150. https://doi.org/10.1016/S0021-9258(18)91404-3

Li, D., Gao, Q., Xu, L., Pang, S., Liu, Z., Wang, C., Tan, W., 2017. Characterization of glutathione S-transferases in the detoxification of metolachlor in two maize cultivars of differing herbicide tolerance. Pestic Biochem Physiol 143, 265–271. https://doi.org/10.1016/J.PESTBP.2016.12.003

Li, F.-W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M.S., Bennett, T., Blázquez, M.A., Cheng, S., Cuming, A.C., de Vries, J., de Vries, S., Delaux, P.-M., Diop, I.S., Harrison, C.J., Hauser, D., Hernández-García, J., Kirbis, A., Meeks, J.C., Monte, I., Mutte, S.K., Neubauer, A., Quandt, D., Robison, T., Shimamura, M., Rensing, S.A., Villarreal, J.C., Weijers, D., Wicke, S., Wong, G.K.-S., Sakakibara, K., Szövényi, P., 2020. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants 6, 259–272. https://doi.org/10.1038/s41477-020-0618-2

Li, J., Mu, J., Bai, J., Fu, F., Zou, T., An, F., Zhang, J., Jing, H., Wang, Q., Li, Z., Yang, S., Zuo, J., 2013. PARAQUAT RESISTANT1, a Golgi-Localized Putative Transporter Protein, Is Involved in Intracellular Transport of Paraquat. Plant Physiol 162, 470–483. https://doi.org/10.1104/pp.113.213892

Li, X., Schuler, M.A., Berenbaum, M.R., 2007. Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. Annu Rev Entomol 52, 231–253. https://doi.org/10.1146/annurev.ento.51.110104.151104

Lichtfouse, E., Navarrete, M., Debaeke, P., Souchère, V., Alberola, C., Ménassieu, J., 2009. Agronomy for Sustainable Agriculture: A Review, in: Sustainable Agriculture. Springer Netherlands, Dordrecht, pp. 1–7. https://doi.org/10.1007/978-90-481-2666-8_1

Ligrone, R., Duckett, J.G., Renzaglia, K.S., 2012. Major transitions in the evolution of early land plants: a bryological perspective. Ann Bot 109, 851–871. https://doi.org/10.1093/aob/mcs017

Lira, C.E., Ahn, J.H., Lim, J., 2006. Molecular Genetic Analysis of Tandemly Located Glycosyltransferase Genes, UGT73B1, UGT73B2, and UGT73B3, in *Arabidopsis thaliana*, Journal of Plant Biology.

Liu, W., Bai, S., Zhao, N., Jia, S., Li, W., Zhang, L., Wang, J., 2018. Non-target site-based resistance to tribenuron-methyl and essential involved genes in *Myosoton aquaticum* (L.). BMC Plant Biol 18, 225. https://doi.org/10.1186/s12870-018-1451-x

Locher, K.P., 2016. Mechanistic diversity in ATP-binding cassette (ABC) transporters. Nat Struct Mol Biol 23, 487–493. https://doi.org/10.1038/nsmb.3216

López-Méndez, B., Baron, B., Brautigam, C.A., Jowitt, T.A., Knauer, S.H., Uebel, S., Williams, M.A., Sedivy, A., Abian, O., Abreu, C., Adamczyk, M., Bal, W., Berger, S., Buell, A.K., Carolis, C., Daviter, T., Fish, A., Garcia-Alai, M., Guenther, C., Hamacek, J., Holková, J., Houser, J., Johnson, C., Kelly, S., Leech, A., Mas, C., Matulis, D., McLaughlin, S.H., Montserret, R., Nasreddine, R., Nehmé, R., Nguyen, Q., Ortega-Alarcón, D., Perez, K., Pirc, K., Piszczek, G., Podobnik, M., Rodrigo, N., Rokov-Plavec, J., Schaefer, S., Sharpe, T., Southall, J., Staunton, D., Tavares, P., Vanek, O., Weyand, M., Wu, D., 2021a. Reproducibility and accuracy of microscale thermophoresis in the NanoTemper Monolith: a multi laboratory benchmark study. Eur Biophys J 50, 411–427. https://doi.org/10.1007/s00249-021-01532-6

López-Méndez, B., Uebel, S., Lundgren, L.P., Sedivy, A., 2021b. Microscale Thermophoresis and additional effects measured in NanoTemper Monolith instruments. European Biophysics Journal 50, 653–660. https://doi.org/10.1007/s00249-021-01529-1

Lorraine-Colwill, D.F., Powles, S.B., Hawkes, T.R., Hollinshead, P.H., Warner, S.A.J., Preston, C., 2002. Investigations into the mechanism of glyphosate resistance in *Lolium rigidum*. Pestic Biochem Physiol 74, 62–72. https://doi.org/10.1016/S0048-3575(03)00007-5

Loutre, C., Dixon, D.P., Brazier, M., Slater, M., Cole, D.J., Edwards, R., 2003. Isolation of a glucosyltransferase from *Arabidopsis thaliana* active in the metabolism of the persistent pollutant 3,4-dichloroaniline. The Plant Journal 34, 485–493. https://doi.org/10.1046/j.1365-313X.2003.01742.x

Love, M.I., Anders, S., Huber, W., 2016. Differential analysis of count data-the DESeq2 package. https://doi.org/10.1186/s13059-014-0550-8

Lu, Y.-P., Li, Z.-S., Rea, P.A., 1997. *AtMRP1* gene of *Arabidopsis* encodes a glutathione *S* -conjugate pump: Isolation and functional definition of a plant ATP-binding cassette transporter gene. Proceedings of the National Academy of Sciences 94, 8243–8248. https://doi.org/10.1073/pnas.94.15.8243

Lundstrom, T.S., Sobel, J.D., 1995. Vancomycin, Trimethoprimsulfamethoxole, and Rifampin. Infect Dis Clin North Am 9, 747–767. https://doi.org/10.1016/S0891-5520(20)30695-4

Lynch, J.H., 2022. Revisiting the dual pathway hypothesis of Chorismate production in plants. Hortic Res 9. https://doi.org/10.1093/hr/uhac052

Machida, M., Takechi, K., Sato, H., Chung, S.J., Kuroiwa, H., Takio, S., Seki, M., Shinozaki, K., Fujita, T., Hasebe, M., Takano, H., 2006. Genes for the peptidoglycan synthesis pathway are essential for chloroplast division in moss. Proceedings of the National Academy of Sciences 103, 6753–6758. https://doi.org/10.1073/pnas.0510693103

Maeda, H., Dudareva, N., 2012. The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. Annu Rev Plant Biol 63, 73–105. https://doi.org/10.1146/annurev-arplant-042811-105439

Marble, S.C., Frank, M.S., Laughinghouse, D.D., Steed, S.T., Boyd, N.S., 2017. Biology and Management of Liverwort (*Marchantia polymorpha*) in Ornamental Crop Production. EDIS 2017. https://doi.org/10.32473/edis-ep542-2017

Maroli, A., Nandula, V., Duke, S., Tharayil, N., 2016. Stable Isotope Resolved Metabolomics Reveals the Role of Anabolic and Catabolic Processes in Glyphosate-Induced Amino

Acid Accumulation in *Amaranthus palmeri* Biotypes. J Agric Food Chem 64, 7040–7048. https://doi.org/10.1021/acs.jafc.6b02196
Martinoia, E., Grill, E., Tommasini, R., Kreuz, K., Amrhein, N., 1993. ATP-dependent glutathione S-conjugate "export" pump in the vacuolar membrane of plants. Nature 364, 247–249. https://doi.org/10.1038/364247a0
Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kolukisaoglu, Ü., Müller-Röber, B., Schulz, B., 2002. Multifunctionality of plant ABC transporters – more than just detoxifiers. Planta 214, 345–355. https://doi.org/10.1007/s004250100661
Matsumoto, H., Takechi, K., Sato, H., Takio, S., Takano, H., 2012. Treatment with Antibiotics that Interfere with Peptidoglycan Biosynthesis Inhibits Chloroplast Division in the Desmid Closterium. PLoS One 7, e40734. https://doi.org/10.1371/journal.pone.0040734
Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, Keiji, Yagisawa, F., Nishida, Keishin, Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara, Y., Kuroiwa, T., 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428, 653–657. https://doi.org/10.1038/nature02398
Matysik., J, Alia, B.B. and P.M., 2002. Molecular mechanisms of quenching of reactive oxygen species by proline under stress in plants on JSTOR. Curr Sci 82, 525–532.
Matzrafi, M., Shaar-Moshe, L., Rubin, B., Peleg, Z., 2017. Unraveling the Transcriptional Basis of Temperature-Dependent Pinoxaden Resistance in *Brachypodium hybridum*. Front Plant Sci 8, 1064. https://doi.org/10.3389/fpls.2017.01064
Meech, R., Mackenzie, P.I., 1997. Structure and function of uridine diphosphate glucuronosyltransferases. Clin Exp Pharmacol Physiol 24, 907–915. https://doi.org/10.1111/j.1440-1681.1997.tb02718.x
Menendez, J., Rojano-Delgado, M.A., De Prado, R., 2014. Differences in Herbicide Uptake, Translocation, and Distribution as Sources of Herbicide Resistance in Weeds, in: ACS Symposium Series. pp. 141–157. https://doi.org/10.1021/bk-2014-1171.ch007
Mentewab, A., Stewart, C.N., 2005. Overexpression of an *Arabidopsis thaliana* ABC transporter confers kanamycin resistance to transgenic plants. Nat Biotechnol 23, 1177–1180. https://doi.org/10.1038/nbt1134
Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Grigoriev, I. V., Rokhsar, D.S., Grossman, A.R., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S. V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J.,

Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W.C.A., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science (1979) 318, 245–251. https://doi.org/10.1126/science.1143609

Mesnage, R., Agapito-Tenfen, S.Z., Vilperte, V., Renney, G., Ward, M., Séralini, G.-E., Nodari, R.O., Antoniou, M.N., 2016. An integrated multi-omics analysis of the NK603 Roundup-tolerant GM maize reveals metabolism disturbances caused by the transformation process. Sci Rep 6, 37855. https://doi.org/10.1038/srep37855

Molin, W.T., Yaguchi, A., Blenner, M., Saski, C.A., 2020. The EccDNA Replicon: A Heritable, Extranuclear Vehicle That Enables Gene Amplification and Glyphosate Resistance in *Amaranthus palmeri*. Plant Cell 32, 2132–2140. https://doi.org/10.1105/tpc.20.00099

Morant, M., Bak, S., Møller, B.L., Werck-Reichhart, D., 2003. Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr Opin Biotechnol 14, 151–162. https://doi.org/10.1016/S0958-1669(03)00024-7

Moretti, M.L., Van Horn, C.R., Robertson, R., Segobye, K., Weller, S.C., Young, B.G., Johnson, W.G., Douglas Sammons, R., Wang, D., Ge, X., d'Avignon, A., Gaines, T.A., Westra, P., Green, A.C., Jeffery, T., Lespérance, M.A., Tardif, F.J., Sikkema, P.H., Christopher Hall, J., McLean, M.D., Lawton, M.B., Schulz, B., 2018. Glyphosate resistance in *Ambrosia trifida:* Part 2. Rapid response physiology and non-target-site resistance. Pest Manag Sci 74, 1079–1088. https://doi.org/10.1002/ps.4569

Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Yang, Z., Schneider, H., Donoghue, P.C.J., 2018. The timescale of early land plant evolution. Proc Natl Acad Sci USA 115, E2274–E2283. https://doi.org/10.1073/pnas.1719588115

Morrison, J.E., Gerik, T.J., Chichester, F.W., Martin, J.R., Chandler, J.M., 1990. A No-Tillage Farming System for Clay Soils. Journal of Production Agriculture 3, 219–227. https://doi.org/10.2134/jpa1990.0219

Moss, S.R., Perryman, S.A.M., Tatnell, L. V., 2007. Managing Herbicide-resistant Blackgrass (*Alopecurus Myosuroides*): Theory and Practice. Weed Technology 21, 300–309. https://doi.org/10.1614/WT-06-087.1

Mougin, C., Polge, N., Scalla, R., Cabanne, F., 1991. Interactions of various agrochemicals with cytochrome P-450-dependent monooxygenases of wheat cells. Pestic Biochem Physiol 40, 1–11. https://doi.org/10.1016/0048-3575(91)90044-M

Murphy, B.P., Tranel, P.J., 2019. Target-Site Mutations Conferring Herbicide Resistance. Plants 8, 382. https://doi.org/10.3390/plants8100382

Nandula, V.K., Riechers, D.E., Ferhatoglu, Y., Barrett, M., Duke, S.O., Dayan, F.E., Goldberg-Cavalleri, A., Tétard-Jones, C., Wortley, D.J., Onkokesung, N., Brazier-Hicks, M., Edwards, R., Gaines, T., Iwakami, S., Jugulam, M., Ma, R., 2019. Herbicide Metabolism: Crop Selectivity, Bioactivation, Weed Resistance, and Regulation. Weed Sci 67, 149–175. https://doi.org/10.1017/wsc.2018.88

Nie, Y., Foster, C.S.P., Zhu, T., Yao, R., Duchêne, D.A., Ho, S.Y.W., Zhong, B., 2020. Accounting for Uncertainty in the Evolutionary Timescale of Green Plants Through Clock-Partitioning and Fossil Calibration Strategies. Syst Biol 69, 1–16. https://doi.org/10.1093/sysbio/syz032

Nikolenko, S.I., Korobeynikov, A.I., Alekseyev, M.A., 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. BMC Genomics 14, S7. https://doi.org/10.1186/1471-2164-14-S1-S7

Nishiyama, T., Wolf, P.G., Kugita, M., Sinclair, R.B., Sugita, M., Sugiura, C., Wakasugi, T., Yamada, K., Yoshinaga, K., Yamaguchi, K., Ueda, K., Hasebe, M., 2004. Chloroplast Phylogeny Indicates that Bryophytes Are Monophyletic. Mol Biol Evol 21, 1813–1819. https://doi.org/10.1093/molbev/msh203

Nol, N., Tsikou, D., Eid, M., Livieratos, I.C., Giannopolitis, C.N., 2012. Shikimate leaf disc assay for early detection of glyphosate resistance in *Conyza canadensis* and relative transcript levels of EPSPS and ABC transporter genes. Weed Res 52, 233–241. https://doi.org/10.1111/j.1365-3180.2012.00911.x

Nuria de María, José M. Becerril, José I. García-Plazaola, Antonio Hernández, María R. de Felipe, and, Mercedes Fernández-Pascual, 2006. New Insights on Glyphosate Mode of Action in Nodular Metabolism: Role of Shikimate Accumulation. https://doi.org/10.1021/JF058166C

Oerke, E.-C., 2006. Crop losses to pests. J Agric Sci 144, 31–43. https://doi.org/10.1017/S0021859605005708

Ohkawa, H., Tsujii, H., Ohkawa, Y., 1999. The use of cytochrome P450 genes to introduce herbicide tolerance in crops: a review. Pestic Sci 55, 867–874. https://doi.org/10.1002/(SICI)1096-9063(199909)55:9<867::AID-PS31>3.0.CO;2-S

Okonechnikov, K., Golosova, O., Fursov, M., team, the U., 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28, 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Orcaray, L., Igal, M., Marino, D., Zabalza, A., Royuela, M., 2010. The possible role of quinate in the mode of action of glyphosate and acetolactate synthase inhibitors. Pest Manag Sci 66, 262–269. https://doi.org/10.1002/ps.1868

Osmani, S.A., Bak, S., Møller, B.L., 2009. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry 70, 325–347. https://doi.org/10.1016/J.PHYTOCHEM.2008.12.009

Padgette, S.R., Kolacz, K.H., Delannay, X., Re, D.B., LaVallee, B.J., Tinius, C.N., Rhodes, W.K., Otero, Y.I., Barry, G.F., Eichholtz, D.A., Peschke, V.M., Nida, D.L., Taylor, N.B., Kishore, G.M., 1995. Development, Identification, and Characterization of a Glyphosate-Tolerant Soybean Line. Crop Sci 35, 1451–1461. https://doi.org/10.2135/cropsci1995.0011183X003500050032x

Padgette, S.R., Re, D.B., Barry, G.F., Eichholtz, D.E., Xavier, Delannay., Fuchs, R.L., Kishore, G.M., Fraley, R.T., 1988. New Weed Control Opportunities: Development of Soybeans With A Roundup Ready™ Gene, in: Herbicide- Resistant Crops. CRC Press, pp. 53–84. https://doi.org/10.1201/9781351073196-4

Pan, L., Gao, H., Xia, W., Zhang, T., Dong, L., 2016. Establishing a herbicide-metabolizing enzyme library in *Beckmannia syzigachne* to identify genes associated with metabolic resistance. J Exp Bot 67, 1745–1757. https://doi.org/10.1093/jxb/erv565

Pan, L., Yu, Q., Han, H., Mao, L., Nyporko, A., Fan, L., Bai, L., Powles, S., 2019. Aldo-keto Reductase Metabolizes Glyphosate and Confers Glyphosate Resistance in *Echinochloa colona*. Plant Physiol 181, 1519–1534. https://doi.org/10.1104/pp.19.00979

Pan, L., Yu, Q., Wang, J., Han, H., Mao, L., Nyporko, A., Maguza, A., Fan, L., Bai, L., Powles, S., 2021. An ABCC-type transporter endowing glyphosate resistance in plants. Proceedings

of the National Academy of Sciences 118, e2100136118. https://doi.org/10.1073/pnas.2100136118
Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197
Perotti, V.E., Larran, A.S., Palmieri, V.E., Martinatto, A.K., Alvarez, C.E., Tuesca, D., Permingeat, H.R., 2019. A novel triple amino acid substitution in the EPSPS found in a high-level glyphosate-resistant *Amaranthus hybridus* population from Argentina. Pest Manag Sci 75, 1242–1251. https://doi.org/10.1002/ps.5303
Peterson, M.A., Collavo, A., Ovejero, R., Shivrain, V., Walsh, M.J., 2018. The challenge of herbicide resistance around the world: a current summary. Pest Manag Sci 74, 2246–2259. https://doi.org/10.1002/ps.4821
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605–12. https://doi.org/10.1002/jcc.20084
Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX : Structure visualization for researchers, educators, and developers. Protein Science 30, 70–82. https://doi.org/10.1002/pro.3943
Phalan, B., Onial, M., Balmford, A., Green, R.E., 2011. Reconciling Food Production and Biodiversity Conservation: Land Sharing and Land Sparing Compared. Science (1979) 333, 1289–1291. https://doi.org/10.1126/science.1208742
Phytozome v12.1: Search for genes, families and sequences [WWW Document], n.d. URL https://phytozome.jgi.doe.gov/pz/portal.html#!search?show=BLAST (accessed 6.24.20).
Piasecki, C., Yang, Y., Benemann, D.P., Kremer, F.S., Galli, V., Millwood, R.J., Cechin, J., Agostinetto, D., Maia, L.C., Vargas, L., Stewart, C.N., 2019. Transcriptomic Analysis Identifies New Non-Target Site Glyphosate-Resistance Genes in *Conyza bonariensis*. Plants 8, 157. https://doi.org/10.3390/plants8060157
Pires, N.D., Dolan, L., 2012. Morphological evolution in land plants: new designs with old genes. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 508–518. https://doi.org/10.1098/rstb.2011.0252
Pollegioni, L., Schonbrunn, E., Siehl, D., 2011. Molecular basis of glyphosate resistance - different approaches through protein engineering. FEBS Journal 278, 2753–2766. https://doi.org/10.1111/j.1742-4658.2011.08214.x
Powles, S.B., Lorraine-Colwill, D.F., Dellow, J.J., Preston, C., 1998. Evolved resistance to glyphosate in rigid ryegrass (*Lolium rigidum*) in Australia. Weed Sci 46, 604–607. https://doi.org/10.1017/S0043174500091165
Powles, S.B., Yu, Q., 2010. Evolution in Action: Plants Resistant to Herbicides. Annu Rev Plant Biol 61, 317–347. https://doi.org/10.1146/annurev-arplant-042809-112119
Pratley, J., Urwin, N., Stanton, R., Baines, P., Broster, J., Cullis, K., Schafer, D., Bohn, J., Krueger, R., 1999. Resistance to glyphosate in *Lolium rigidum.* I. Bioevaluation. Weed Sci 47, 405–411. https://doi.org/10.1017/S0043174500091992
Price, D.C., Goodenough, U.W., Roth, R., Lee, J.-H., Kariyawasam, T., Mutwil, M., Ferrari, C., Facchinelli, F., Ball, S.G., Cenci, U., Chan, C.X., Wagner, N.E., Yoon, H.S., Weber, A.P.M., Bhattacharya, D., 2019. Analysis of an improved *Cyanophora paradoxa* genome assembly. DNA Res 26, 287–299. https://doi.org/10.1093/dnares/dsz009

Priestman, M.A., Healy, M.L., Funke, T., Becker, A., Schönbrunn, E., 2005. Molecular basis for the glyphosate-insensitivity of the reaction of 5-enolpyruvylshikimate 3-phosphate synthase with shikimate. FEBS Lett 579, 5773–5780. https://doi.org/10.1016/J.FEBSLET.2005.09.066

Queiroz, A.R.S. de, Delatorre, C.A., Lucio, F.R., Rossi, C.V.S., Zobiole, L.H.S., Merotto, A., 2019. Rapid necrosis: a novel plant resistance mechanism to 2,4-D. Weed Sci 68, 1–13. https://doi.org/10.1017/wsc.2019.65

Rambaut, A., 2007. FigTree [WWW Document]. URL http://tree.bio.ed.ac.uk/software/figtree/ (accessed 5.30.23).

Rangani, G., Rouse, C.E., Saski, C., Noorai, R.E., Shankar, V., Lawton-Rauh, A.L., Werle, I.S., Roma-Burgos, N., 2022. High Resistance to Quinclorac in Multiple-Resistant *Echinochloa colona* Associated with Elevated Stress Tolerance Gene Expression and Enriched Xenobiotic Detoxification Pathway. Genes (Basel) 13, 515. https://doi.org/10.3390/genes13030515

Ravinder S. Sidhu, Bruce G. Hammond, Roy L. Fuchs, Jean-Noel Mutz, Larry R. Holden, Beverly George, and, Olson, T., 2000. Glyphosate-Tolerant Corn: The Composition and Feeding Value of Grain from Glyphosate-Tolerant Corn Is Equivalent to That of Conventional Corn (*Zea mays L.*). https://doi.org/10.1021/JF000172F

Reade, J.P.H., Milner, L.J., Cobb, A.H., 2004. A role for glutathione *S* -transferases in resistance to herbicides in grasses. Weed Sci 52, 468–474. https://doi.org/10.1614/P2002-168D

Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S.I., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W.B., Barker, E., Bennetzen, J.L., Blankenship, R., Sung, H.C., Dutcher, S.K., Estelle, M., Fawcett, J.A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K.A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D.R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P.J., Sanderfoot, A., Schween, G., Shiu, S.H., Stueber, K., Theodoulou, F.L., Tu, H., Van De Peer, Y., Verrier, P.J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A.C., Hasebe, M., Lucas, S., Mishler, B.D., Reski, R., Grigoriev, I. V., Quatrano, R.S., Boore, J.L., 2008. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science (1979) 319, 64–69. https://doi.org/10.1126/science.1150646

Rice Genome Annotation Project [WWW Document], n.d. URL http://rice.plantbiology.msu.edu/ (accessed 6.24.20).

Rodríguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., Lang, B.F., 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15, 1325–30. https://doi.org/10.1016/j.cub.2005.06.040

Ross, J., Li, Y., Lim, E.-K., Bowles, D.J., 2001. Higher plant glycosyltransferases. Genome Biol 2, reviews3004.1. https://doi.org/10.1186/gb-2001-2-2-reviews3004

Russell Thom, Ian Cummins, David P. Dixon, Robert Edwards, David J. Cole, and Adrian J. Lapthorn, 2002. Structure of a Tau Class Glutathione S-Transferase from Wheat Active in Herbicide Detoxification. https://doi.org/10.1021/BI015964X

Saha, J., Sengupta, A., Gupta, K., Gupta, B., 2015. Molecular phylogenetic study and expression analysis of ATP-binding cassette transporter gene family in *Oryza sativa* in

response to salt stress. Comput Biol Chem 54, 18–32. https://doi.org/10.1016/J.COMPBIOLCHEM.2014.11.005

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., Toki, S., 2014. A novel rice cytochrome P450 gene, CYP72A31, confers tolerance to acetolactate synthase-inhibiting herbicides in rice and Arabidopsis. Plant Physiol 166, 1232–40. https://doi.org/10.1104/pp.113.231266

Sakiko Hirose, Hiroyuki Kawahigashi, Kenjirou Ozawa, Noriaki Shiota, Hideyuki Inui, Hideo Ohkawa, and Ohkawa, Y., 2005. Transgenic Rice Containing Human CYP2B6 Detoxifies Various Classes of Herbicides. https://doi.org/10.1021/JF050064Z

Salas, R.A., Dayan, F.E., Pan, Z., Watson, S.B., Dickson, J.W., Scott, R.C., Burgos, N.R., 2012. EPSPS gene amplification in glyphosate-resistant Italian ryegrass (*Lolium perenne ssp*. *multiflorum*) from Arkansas. Pest Manag Sci 68, 1223–1230. https://doi.org/10.1002/ps.3342

Salas-Perez, R.A., Saski, C.A., Noorai, R.E., Srivastava, S.K., Lawton-Rauh, A.L., Nichols, R.L., Roma-Burgos, N., 2018. RNA-Seq transcriptome analysis of *Amaranthus palmeri* with differential tolerance to glufosinate herbicide. PLoS One 13, 1–33. https://doi.org/10.1371/journal.pone.0195488

Sammons, R.D., Gaines, T.A., 2014. Glyphosate resistance: state of knowledge. Pest Manag Sci 70, 1367–1377. https://doi.org/10.1002/ps.3743

Sato, N., Toyoshima, M., Tajima, N., Takechi, K., Takano, H., 2017. Single-Pixel Densitometry Revealed the Presence of Peptidoglycan in the Intermembrane Space of the Moss Chloroplast Envelope in Conventional Electron Micrographs. Plant Cell Physiol 58, 1743–1751. https://doi.org/10.1093/pcp/pcx113

Sauret-Güeto, S., Frangedakis, E., Silvestri, L., Rebmann, M., Tomaselli, M., Markel, K., Delmans, M., West, A., Patron, N.J., Haseloff, J., 2020. Systematic Tools for Reprogramming Plant Gene Expression in a Simple Model, *Marchantia polymorpha*. ACS Synth Biol 9, 864–882. https://doi.org/10.1021/acssynbio.9b00511

Scheuermann, Thomas H., Padrick, S.B., Gardner, K.H., Brautigam, C.A., 2016. On the acquisition and analysis of microscale thermophoresis data. Anal Biochem 496, 79–93. https://doi.org/10.1016/J.AB.2015.12.013

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Schönbrunn, E., Eschenburg, S., Shuttleworth, W.A., Schloss, J. V., Amrhein, N., Evans, J.N.S., Kabsch, W., 2001. Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. Proc Natl Acad Sci U S A 98, 1376. https://doi.org/10.1073/PNAS.98.4.1376

Schuler, M.A., 1996. Plant Cytochrome P450 Monooxygenases. CRC Crit Rev Plant Sci 15, 235–284. https://doi.org/10.1080/07352689609701942

Schulz, B., Kolukisaoglu, H.U., 2006. Genomics of plant ABC transporters: the alphabet of photosynthetic life forms or just holes in membranes? FEBS Lett 580, 1010–6. https://doi.org/10.1016/j.febslet.2006.01.002

Sen, M.K., Hamouzová, K., Mikulka, J., Bharati, R., Košnarová, P., Hamouz, P., Roy, A., Soukup, J., 2021. Enhanced metabolism and target gene overexpression confer resistance

against acetolactate synthase-inhibiting herbicides in *Bromus sterilis*. Pest Manag Sci 77, 2122–2128. https://doi.org/10.1002/ps.6241
Shah, D.M., Hironaka, C.M., Wiegand, R.C., Harding, E.I., Krivi, G.G., Tiemeier, D.C., 1986. Structural analysis of a maize gene coding for glutathione-S-transferase involved in herbicide detoxification, Plant Molecular Biology.
Shimamura, M., 2016. *Marchantia polymorpha* : Taxonomy, Phylogeny and Morphology of a Model System. Plant Cell Physiol 57, 230–256. https://doi.org/10.1093/pcp/pcv192
Siehl, D.L., 1997. Inhibitors of EPSP Synthase, Glutamine Synthetase and Histidine Synthesis. 105 Press.
Siminszky, B., 2006. Plant cytochrome P450-mediated herbicide metabolism. Phytochemistry Reviews 5, 445–458. https://doi.org/10.1007/s11101-006-9011-7
Skarzynski, T., Mistry, A., Wonacott, A., Hutchinson, S.E., Kelly, V.A., Duncan, K., 1996. Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-N-acetylglucosamine and the drug fosfomycin. Structure 4, 1465–74. https://doi.org/10.1016/s0969-2126(96)00153-0
Sousa, F., Foster, P.G., Donoghue, P.C.J., Schneider, H., Cox, C.J., 2019. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). New Phytologist 222, 565–575. https://doi.org/10.1111/nph.15587
Su, D., Yang, L., Shi, X., Ma, X., Zhou, X., Hedges, S.B., Zhong, B., 2021. Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants. Mol Biol Evol 38, 3332–3344. https://doi.org/10.1093/molbev/msab106
Sugano, S.S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I., Kohchi, T., 2014. CRISPR/Cas9-mediated targeted mutagenesis in the liverwort Marchantia polymorpha L. Plant Cell Physiol 55, 475–81. https://doi.org/10.1093/pcp/pcu014
Sweetser, P.B., Schow, G.S., Hutchison, J.M., 1982. Metabolism of chlorsulfuron by plants: Biological basis for selectivity of a new herbicide for cereals. Pestic Biochem Physiol 17, 18–23. https://doi.org/10.1016/0048-3575(82)90121-3
Takano, H., Tsunefuka, T., Takio, S., Ishikawa, H., Cytologia, K.T.-, 2018. Visualization of plastid peptidoglycan in the charophyte alga *Klebsormidium nitens* using a metabolic labeling method. jstage.jst.go.jp.
Takano, H.K., Patterson, E.L., Nissen, S.J., Dayan, F.E., Gaines, T.A., 2019. Predicting herbicide movement across semi-permeable membranes using three phase partitioning. Pestic Biochem Physiol 159, 22–26. https://doi.org/10.1016/J.PESTBP.2019.05.009
Tani, E., Chachalis, D., Travlos, I., Bilalis, D., 2016. Environmental Conditions Influence Induction of Key ABC-Transporter Genes Affecting Glyphosate Resistance Mechanism in *Conyza canadensis*. Int J Mol Sci 17, 342. https://doi.org/10.3390/ijms17040342
The Arabidopsis Information Resource (TAIR) [WWW Document], n.d. URL https://www.arabidopsis.org/Blast/index.jsp
Theodoulou, F.L., 2000. Plant ABC transporters. Biochimica et Biophysica Acta (BBA) - Biomembranes 1465, 79–103. https://doi.org/10.1016/S0005-2736(00)00132-2
Theodoulou, F.L., Kerr, I.D., 2015. ABC transporter research: going strong 40 years on. Biochem Soc Trans 43, 1033–1040. https://doi.org/10.1042/BST20150139
Tiburcio, A.F., Altabella, T., Bitrián, M., Alcázar, R., 2014. The roles of polyamines during the lifespan of plants: from development to stress. Planta 240, 1–18. https://doi.org/10.1007/s00425-014-2055-9

Tiwari, P., Sangwan, R.S., Sangwan, N.S., 2016. Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. Biotechnol Adv 34, 714–739. https://doi.org/10.1016/J.BIOTECHADV.2016.03.006

Tounou, E., Takio, S., Sakai, A., Ono, K., Takano, H., 2002. Ampicillin Inhibits Chloroplast Division in Cultured Cells of the Liverwort *Marchantia polymorpha*. Cytologia (Tokyo) 67, 429–434.

Treutter, D., 2006. Significance of flavonoids in plant resistance: a review. Environ Chem Lett 4, 147–157. https://doi.org/10.1007/s10311-006-0068-8

Tso, S.-C., Chen, Q., Vishnivetskiy, S.A., Gurevich, V. V., Iverson, T.M., Brautigam, C.A., 2018. Using two-site binding models to analyze microscale thermophoresis data. Anal Biochem 540–541, 64–75. https://doi.org/10.1016/J.AB.2017.10.013

Tzin, V., Galili, G., 2010. The Biosynthetic Pathways for Shikimate and Aromatic Amino Acids in *Arabidopsis thaliana*. Arabidopsis Book 8, e0132. https://doi.org/10.1199/tab.0132

Usui, K., 2001. Metabolism and selectivity of rice herbicides in plants. Weed Biol Manag 1, 137–146. https://doi.org/10.1046/j.1445-6664.2001.00029.x

Utsunomiya, H., Saiki, N., Kadoguchi, H., Fukudome, M., Hashimoto, S., Ueda, M., Takechi, K., Takano, H., 2021. Genes encoding lipid II flippase MurJ and peptidoglycan hydrolases are required for chloroplast division in the moss *Physcomitrella patens*. Plant Mol Biol 107, 405–415. https://doi.org/10.1007/s11103-020-01081-0

Van Horn, C.R., Moretti, M.L., Robertson, R.R., Segobye, K., Weller, S.C., Young, B.G., Johnson, W.G., Schulz, B., Green, A.C., Jeffery, T., Lespérance, M.A., Tardif, F.J., Sikkema, P.H., Hall, J.C., McLean, M.D., Lawton, M.B., Sammons, R.D., Wang, D., Westra, P., Gaines, T.A., 2018. Glyphosate resistance in *Ambrosia trifida*: Part 1. Novel rapid cell death response to glyphosate. Pest Manag Sci 74, 1071–1078. https://doi.org/10.1002/ps.4567

Verrier, P.J., Bird, D., Burla, B., Dassa, E., Forestier, C., Geisler, M., Klein, M., Kolukisaoglu, Ü., Lee, Y., Martinoia, E., Murphy, A., Rea, P.A., Samuels, L., Schulz, B., Spalding, E.P., Yazaki, K., Theodoulou, F.L., 2008. Plant ABC proteins – a unified nomenclature and updated inventory. Trends Plant Sci 13, 151–159. https://doi.org/10.1016/J.TPLANTS.2008.02.001

Vila-Aiub, M.M., Goh, S.S., Gaines, T.A., Han, H., Busi, R., Yu, Q., Powles, S.B., 2014. No fitness cost of glyphosate resistance endowed by massive EPSPS gene amplification in *Amaranthus palmeri*. Planta 239, 793–801. https://doi.org/10.1007/s00425-013-2022-x

Wakelin, A.M., Lorraine-Colwill, D.F., Preston, C., 2004. Glyphosate resistance in four different populations of *Lolium rigidum* is associated with reduced translocation of glyphosate to meristematic zones. Weed Res 44, 453–459. https://doi.org/10.1111/j.1365-3180.2004.00421.x

Ward, S.M., Webster, T.M., Steckel, L.E., 2013. Palmer Amaranth (*Amaranthus palmeri*): A Review. Weed Technology 27, 12–27. https://doi.org/10.1614/WT-D-12-00113.1

Whatley, J.M., Whatley, F.R., 1981. Chloroplast Evolution, Source: The New Phytologist.

Windsor, B., Roux, S.J., Lloyd, A., 2003. Multiherbicide tolerance conferred by AtPgp1 and apyrase overexpression in *Arabidopsis thaliana*. Nat Biotechnol 21, 428–433. https://doi.org/10.1038/nbt809

WSSA [WWW Document], 1998. URL https://wssa.net/wssa/weed/resistance/herbicide-resistance-and-herbicide-tolerance-definitions/#:~:text=The%20definitions%20for%20these%20terms,lethal%20to%20the%20wild%20type.

Xi, J., Xu, P., Xiang, C.-B., 2012. Loss of AtPDR11, a plasma membrane-localized ABC transporter, confers paraquat tolerance in *Arabidopsis thaliana*. The Plant Journal 69, 782–791. https://doi.org/10.1111/j.1365-313X.2011.04830.x

Xia, J., Psychogios, N., Young, N., Wishart, D.S., 2009. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic Acids Res 37, W652–W660. https://doi.org/10.1093/nar/gkp356

Xiang, W., Wang, X., Ren, T., 2005. Expression of a wheat cytochrome P450 monooxygenase cDNA in yeast catalyzes the metabolism of sulfonylurea herbicides. Pestic Biochem Physiol 85, 1–6. https://doi.org/10.1016/j.pestbp.2005.09.001

Xiong, J., Feng, J., Yuan, D., Zhou, J., Miao, W., 2015. Tracing the structural evolution of eukaryotic ATP binding cassette transporter superfamily. Sci Rep 5, 16724. https://doi.org/10.1038/srep16724

Xu, J., Wang, X., Guo, W., 2015. The cytochrome P450 superfamily: Key players in plant development and defense. J Integr Agric 14, 1673–1686. https://doi.org/10.1016/S2095-3119(14)60980-1

Yang, Q., Deng, W., Li, X., Yu, Q., Bai, L., Zheng, M., 2016. Target-site and non-target-site based resistance to the herbicide tribenuron-methyl in flixweed (*Descurainia sophia L*.). BMC Genomics 17, 551. https://doi.org/10.1186/s12864-016-2915-8

Yanniccari, M., Palma-Bautista, C., Vázquez-García, J.G., Gigón, R., Mallory-Smith, C.A., De Prado, R., 2023. Constitutive overexpression of *EPSPS* by gene duplication is involved in glyphosate resistance in *Salsola tragus*. Pest Manag Sci 79, 1062–1068. https://doi.org/10.1002/ps.7272

Yonekura-Sakakibara, K., Hanada, K., 2011. An evolutionary view of functional diversity in family 1 glycosyltransferases. The Plant Journal 66, 182–193. https://doi.org/10.1111/j.1365-313X.2011.04493.x

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., Bhattacharya, D., 2004. A Molecular Timeline for the Origin of Photosynthetic Eukaryotes. Mol Biol Evol 21, 809–818. https://doi.org/10.1093/molbev/msh075

Yu, J., Hu, F., Dossa, K., Wang, Z., Ke, T., 2017. Genome-wide analysis of UDP-glycosyltransferase super family in *Brassica rapa* and *Brassica oleracea* reveals its evolutionary history and functional characterization. BMC Genomics 18, 474. https://doi.org/10.1186/s12864-017-3844-x

Yu, Q., Powles, S., 2014. Metabolism-based herbicide resistance and cross-resistance in crop weeds: a threat to herbicide sustainability and global crop production. Plant Physiol 166, 1106–18. https://doi.org/10.1104/pp.114.242750

Yuan, J.S., Tranel, P.J., Stewart, C.N., 2007. Non-target-site herbicide resistance: a family business. Trends Plant Sci 12, 6–13. https://doi.org/10.1016/J.TPLANTS.2006.11.001

Yun, M.-S., Yogo, Y., Miura, R., Yamasue, Y., Fischer, A.J., 2005. Cytochrome P-450 monooxygenase activity in herbicide-resistant and -susceptible late watergrass (*Echinochloa phyllopogon*). Pestic Biochem Physiol 83, 107–114. https://doi.org/10.1016/J.PESTBP.2005.04.002

Zhang, D., Li, X., Bei, F., Jin, T., Jia, S., Bu, R., Wang, J., Wang, H., Liu, W., 2022. Investigating the Metabolic Mesosulfuron-Methyl Resistance in *Aegilops tauschii* Coss. By Transcriptome Sequencing Combined with the Reference Genome. J Agric Food Chem 70, 11429–11440. https://doi.org/10.1021/acs.jafc.2c04529

Zhang, J., Fu, X.-X., Li, R.-Q., Zhao, X., Liu, Y., Li, M.-H., Zwaenepoel, A., Ma, H., Goffinet, B., Guan, Y.-L., Xue, J.-Y., Liao, Y.-Y., Wang, Q.-F., Wang, Q.-H., Wang, J.-Y., Zhang, G.-Q.,

Wang, Z.-W., Jia, Y., Wang, M.-Z., Dong, S.-S., Yang, J.-F., Jiao, Y.-N., Guo, Y.-L., Kong, H.-Z., Lu, A.-M., Yang, H.-M., Zhang, S.-Z., Van de Peer, Y., Liu, Z.-J., Chen, Z.-D., 2020. The hornwort genome and early land plant evolution. Nat Plants 6, 107–118. https://doi.org/10.1038/s41477-019-0588-4

Zhang, P., Zhang, Z., Zhang, L., Wang, J., Wu, C., 2020. Glycosyltransferase GT1 family: Phylogenetic distribution, substrates coverage, and representative structural features. Comput Struct Biotechnol J 18, 1383. https://doi.org/10.1016/J.CSBJ.2020.06.003

Zhao, N., Li, W., Bai, S., Guo, W., Yuan, G., Wang, F., Liu, W., Wang, J., 2017. Transcriptome Profiling to Identify Genes Involved in Mesosulfuron-Methyl Resistance in Alopecurus aequalis. Front Plant Sci 8, 1391. https://doi.org/10.3389/fpls.2017.01391

Zhao, Y., Li, W., Sun, L., Wu, R., Xu, H., Su, W., Lu, C., 2023. Candidate Genes Involved in Tolerance to Fenoxaprop-P-Ethyl in Rice Induced by Isoxadifen-Ethyl Hydrolysate. Agronomy 13, 225. https://doi.org/10.3390/agronomy13010225

Zhu, J., Patzoldt, W.L., Shealy, R.T., Vodkin, L.O., Clough, S.J., Tranel, P.J., 2008a. Transcriptome Response to Glyphosate in Sensitive and Resistant Soybean. J Agric Food Chem 56, 6355–6363. https://doi.org/10.1021/jf801254e

Zulet, A., Zabalza, A., Royuela, M., 2013. Phytotoxic and Metabolic Effects of Exogenous Quinate on *Pisum sativum L*. J Plant Growth Regul 32, 779–788. https://doi.org/10.1007/s00344-013-9345-5

Zulet-González, A., Barco-Antoñanzas, M., Gil-Monreal, M., Royuela, M., Zabalza, A., 2020. Increased Glyphosate-Induced Gene Expression in the Shikimate Pathway Is Abolished in the Presence of Aromatic Amino Acids and Mimicked by Shikimate. Front Plant Sci 11, 527853. https://doi.org/10.3389/fpls.2020.00459