Rewinding the tape: experimental evolution of resistance to glyphosate

Submitted by

Lieselot Nguyen

To the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences, January 2025.

Author’s declaration

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

I acknowledge the use of:

- Grammarly (https://www.grammarly.com/) to improve grammar within my thesis.
- Chatgpt (https://chatgpt.com/) to refine English academic writing within my thesis and troubleshoot R and Bash code.

Abstract

Modern agriculture relies on chemical weed control to secure crop yields, but extensive and prolonged herbicide use has imposed strong selective pressures, driving the evolution of resistance in weeds and contaminating non-target ecosystems. Understanding the evolutionary dynamics and outcomes of herbicide selection is critical for addressing these challenges. Microbial model organisms, such as the green alga *Chlamydomonas reinhardtii*, provide an ideal system to study resistance evolution in a controlled laboratory setting, and represents an important group of non-target organisms exposed to herbicide pollution.

This thesis examines the effects of glyphosate selection on adaptation in *C. reinhardtii.* Specifically, the study focused on the evolutionary dynamics, fitness, and genomic changes associated with exposure of *C. reinhardtii* to this herbicide, linking resistance phenotypes to genomic changes under contrasting selection regimes. While most research has focused on sudden exposure to high herbicide doses, gradual dose increases can occur due to the accumulation of herbicides in the environment, with potentially distinct evolutionary impacts. Using experimental evolution and whole genome sequencing, adaptive responses were investigated under two ecological scenarios. The first experiment involved a simple scenario with a single rapid glyphosate dose increase, while the second tested a complex scenario with variable changes in dose rates (rapid, intermediate, and slow rates of change).

The findings from these experiments revealed that rapid dose increases augmented variability and delayed resistance evolution, whereas gradual changes imposed fitness costs. Genome-wide variant detection indicated that glyphosate resistance was not associated with mutations or copy number variation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). Instead, evidence suggests that glyphosate resistance in *C. reinhardtii* is driven by non-target-site mechanisms, regardless of selection history. Furthermore, limited overlap in resistance-associated loci between treatments and replicates suggests a complex, and potentially independent, evolutionary response to glyphosate selection. These insights contribute to our understanding of the evolutionary and genomic mechanisms driving herbicide resistance.

Covid-19 statement

Covid-19 pandemic restrictions occurred during the experimental work described in Chapter 4 and prevented the completion of the full experimental plan. Chapter 4 originally set out to investigate the effect of different selection histories on evolution of glyphosate resistance and associated relative fitness in both the selective environment and ancestral media, at different time points throughout the selection experiment. A follow-up experiment was also planned to determine the levels of glyphosate resistance attained. While the selection experiment and assessing fitness in ancestral and selective environment for populations exposed to the full 12 weeks of selection were carried out in time, the Covid-19 pandemic restrictions prevented access to facilities before populations exposed to 4 and 8 weeks of selection could be assessed. All populations were thus stored on agar slopes with the aim to revive them at a later time, to carry out the fitness assays as well as run a glyphosate dose response experiment to test the effect of selective histories on levels of glyphosate resistance. Conducting fitness assays on populations exposed to 4 and 8 weeks of selection might have helped to uncover when glyphosate resistance arose, and test if it was due to acclimation or adaptation. In the case of selection at different glyphosate doses, these assays would have informed if pre-exposure to lower doses already conferred resistance to 1 MIC, or if adaptation to the lower dose conferred only partial resistance, requiring additional adaptation to gain full resistance. Alternatively, it might have revealed if populations selected under lower dose are acclimated and maintain populations large enough for adaptive mutations to arise once exposed to 1MIC. Unfortunately, when I was allowed access to the laboratory again, more than 6 months had elapsed, and populations stored on agar slopes were lost.

Additionally, loss of cell lines prevented me from carrying out the experiment originally planned for Chapter 5 on relaxed selection. The aim was to investigate the effect of selective histories from Chapter 4 (glyphosate resistance and associated fitness costs evolved under different glyphosate doses), on subsequent loss of resistance in the absence of glyphosate selection. This would have tested the hypotheses that loss of resistance would only occur when associated to a fitness cost.

As a further part of this work, I had planned to use the stored cells from Chapter 4 to investigate the effect of the glyphosate selective environment on the degree of generality evolved. In Chapter 4, Chlamydomonas exposed to both rapid and gradual increases in glyphosate dose were able to evolve glyphosate resistance. However, based upon the higher plant resistance literature, I hypothesised that lower doses, and more gradual increase in selective pressure may have selected for mutations conferring 'generalist' resistance phenotypes. I had planned to take the evolved cell lines generated in Chapter 4, and further screen them for any alteration in sensitivity to a range of other herbicide stressors. If the lower dose / slower selection was contributing to a more generalist resistance phenotype, I would have expected to observe greater cell growth in the presence of these novel selective herbicides for the lines evolved under slower selection, than those exposed directly to 1 MIC of glyphosate.

Since all biological material was lost and whole genome sequencing data was available from data Chapters 3 and 4, this PhD focus turned to uncovering the genetic basis of glyphosate adaptation and the effect of selective histories on selected resistance mechanisms as presented in Chapters 3 and 5.

Acknowledgements

This PhD journey has been a deeply transformative experience, and it would not have been possible without the support, guidance, and encouragement of many remarkable individuals and institutions.

First and foremost, I would like to express my sincere gratitude to my supervisors, Steve Hanley, David Comont, Nick Colegrave and Angus Buckling, for their invaluable guidance, continuous support, and trust in my abilities. Your expertise and insightful feedback have been fundamental in shaping this research, and your patience and encouragement have kept me motivated through the most challenging times.

I am also deeply thankful to Paul Neve, for his constructive criticism, thoughtful suggestions, and encouragement throughout this journey. The insights and perspectives you provided have enriched this work in countless ways.

I extend my appreciation to my colleagues and friends for the stimulating discussions, and camaraderie that made this academic environment a truly inspiring place to work. A special appreciation for Donna Fellowes and Sam Benson, your unwavering friendship and humour have provided much-needed support. I am especially grateful to David Hughes for providing bioinformatics training and support, Rob Ness for sharing advice and datasets for me to train on and Sadik Abrare for providing help with the reference genome annotation. I would also like to thank Suzanne Clark, Kirsty Hassall, Jessy Hood for their guidance through statistical analyses. And all readers and I will be thankful to Sue Steele for proof reading!

I would like to acknowledge the financial support provided by BBSRC and NERC for sequencing which made this research possible.

On a personal note, I owe an enormous debt of gratitude to my family and their belief in me. Special thanks to my partner, Nico, thank you for standing by me with love, understanding, and unwavering support throughout this demanding journey.

Finally, I wish to dedicate this thesis to my daughter, Louenn, which from her canonical 2 years of wisdom kept me grounded and sane in her own special way.

Thank you all for being part of this journey.

List of contents

List of figures

List of tables

List of Appendices

Appendices are in a separate PDF file:

Appendix 1 Glyphosate resistance mechanisms in weeds
Appendix 2: C.reinhardtii DNA extractions protocol
Appendix 3: List of variants susceptible to support glyphosate resistance from the SR experiment (Chapter 3)
Appendix 4: Annotation of Variants susceptible to support glyphosate resistance from the SR experiment (Chapter3)
Appendix 5: List of variants susceptible to support glyphosate resistance from the VR experiment (Chapter5)
Appendix 6: Annotation of Variants susceptible to support glyphosate resistance from the VR experiment (Chapter5)
Appendix 7: PDF version of Figure 30 (Molecular function GO terms heatmap for both SR and VR experiments)

1 General introduction

1.1 Herbicide resistance: the particular case of Glyphosate

1.1.1 Herbicide resistance, an example of evolution in action

Herbicides are chemical compounds used to control unwanted plants referred to as weeds. They are extensively used in agriculture since their first introduction in the 1940's leading to increasing crop yields (Shaner, 2014). However, the use of herbicides also applies a strong selective pressure for weeds to evolve resistance. As modern agriculture has relied mainly on herbicides for weed control, resistance to herbicides has now evolved in a large range of species globally. In some cases, weed populations have evolved resistance to multiple herbicides (Cocker et al., 1999; Tardif and Powles, 1994), and herbicide resistance is now a major threat to food security in the current agricultural model (Oerke,2006). The emergence of herbicide resistance has allowed the scientific community to observe the result of years of human-driven evolution in action (Carroll et al., 2007; Hairston et al., 2005).

1.1.2 Glyphosate use and evolution of resistance overtime

Glyphosate's herbicidal activity was discovered in 1970 by the chemist J.E Frantz who formulated it as the product 'Roundup', and commercialized by Monsanto in 1974 (Benbrook, 2016). Glyphosate [N-(phosphonomethyl) glycine] presented attractive properties: cheap production and highly effective against a broad range of weeds (Baylis, 2000). This non-selective active compound was mainly used as a foliar spray to remove weeds before crop emergence, before the introduction of glyphosate transgenic crops in 1996. Glyphosate resistant (GR) crops allowed multiple glyphosate applications during a cropping season (Richmond, 2018). Monsanto's patent expired in 2000 and other compagnies started manufacturing glyphosate-based herbicides (Richmond, 2018), contributing further to a reduction in glyphosate's cost, and its increased use (Duke and Powles, 2008).

Evolution of glyphosate resistance in weeds was initially considered to be unlikely (Bradshaw et al., 1997). Additionally, glyphosate was thought to have no adverse effects on both human health (Baylis, 2000; Powles, 2008; Williams et al., 2000) and the environment (Giesy et al., 2000) when used according to instructions. Indeed, glyphosate easily binds to soil where microorganism can degrade it into another biodegradable and soil-binding product: aminomethylphosphonic acid (AMPA). Consequently, it was thought that glyphosate and AMPA would not diffuse in the environment. Furthermore, the use of glyphosate to control weeds allowed farmers to practice direct drilling on uncultivated soil, a method saving on fuel and preventing soil erosion and improving its structure (Baylis, 2000; Powles, 2008). Better, cheaper and easier weed management, glyphosate presented all the characteristics of the "perfect herbicide" at the time, which resulted in rapid and widespread adoption of the GR crop technology (Duke and Powles, 2008).

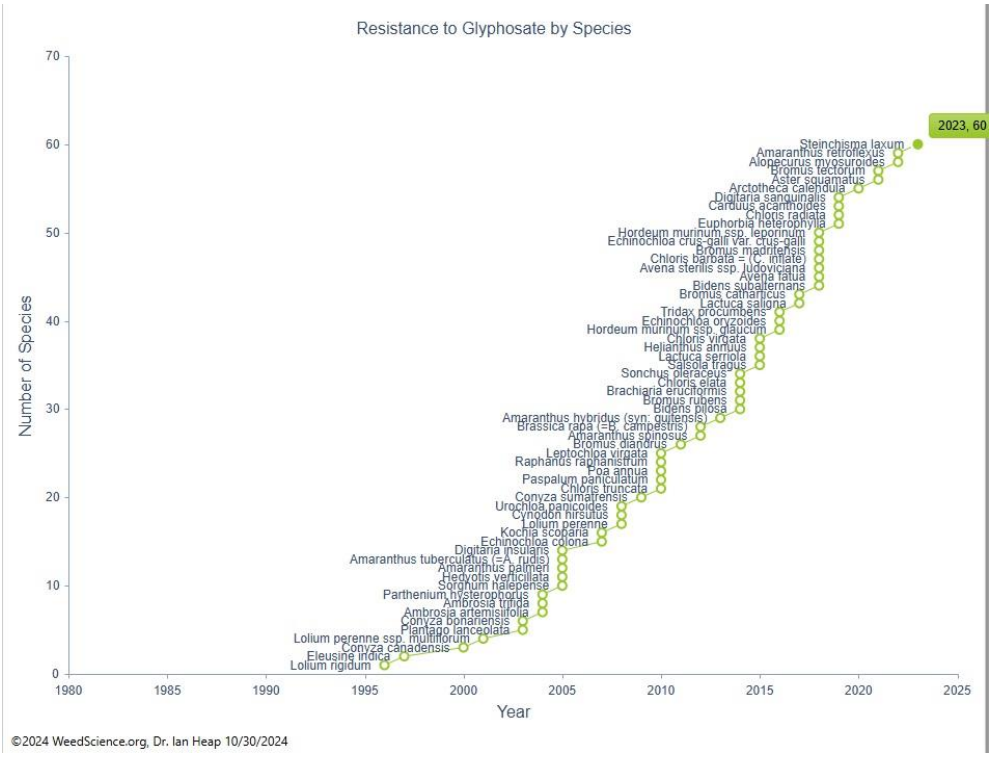

Figure 1: Number of weed species with populations resistant to glyphosate. Source from https://www.weedscience.org/ based on data collected by the Herbicide Resistance Action Committee. Accessed the 10/30/2024.

Similarly to antibiotics, while glyphosate was once seen as a breakthrough that would eradicate weeds when originally introduced, resistance has now widely evolved, with further evidence suggesting that sustained use of glyphosate may also be harmful to the environment and human health (Richmond, 2018). Glyphosate and AMPA are now globally detected in freshwater (Battaglin et al., 2014; Huntscha et al., 2018; Sanchís et al., 2012) and seawater (Wirth et al., 2021). Although present at low concentrations, the ubiquity and potential environmental impact of glyphosate and AMPA is starting to raise concerns (Benbrook, 2016; Gonçalves et al., 2020; Myers et al., 2016; Richmond, 2018). The International Agency for Research on Cancer classified glyphosate as a probable carcinogenic to humans in 2015 drawing closer attention to potential health risks (IARC, 2015; Myers et al., 2016; Richmond, 2018). Additionally, a review in 2008, Duke and Powels warned that the extent of glyphosate's current usage threatened long term efficacy to control weeds. Overreliance on glyphosate has already applied a high selective pressure, leading both to species shifts in weed communities towards naturally resistant species, and evolution of glyphosate resistant weeds.

To mitigate further evolution of glyphosate resistance, Duke and Powels (2008) called to diversify weed management practices in order to keep this cropping system sustainable. Indeed, once herbicide resistance is established and detected in populations it is difficult to offset. Pro-active approaches have therefore been recommended to delay evolution of resistance (Beckie, 2011) , providing long-term cost-benefits in comparison with the usual reactive approach (Edwards et al., 2014). To date more than 60 weed species have independently evolved glyphosate resistance (Figure

1, Heap, 2021), with new resistance cases regularly identified. It is likely that further species have the potential to evolve glyphosate resistance. For example, a recent epidemiological study provided evidence that *Alopecurus myosuroides* UK populations exhibit signatures of selection towards glyphosate resistance, associated with sustained exposure in the field (Comont et al., 2019a).

In summary, if early evidence suggested that the evolution of glyphosate resistant weeds would be unlikely (Bradshaw et al., 1997), this was disproved 24 years after glyphosate's introduction with the first report of a GR weed (Powles et al., 1998).Broad adoption of GR cropping systems and global increases in volumes of glyphosate applied by 15-fold (Benbrook, 2016) have created a world-wide selection experiment (Gaines et al., 2019). Under such a strong selection pressure, a variety of weed species evolved glyphosate resistance based on diverse and sometimes novel resistance mechanisms. To date, glyphosate is the herbicide with the most mechanisms of evolved resistance (Duke et al., 2021). This could be explained by the strong selection pressure and the fact that glyphosate resistance is often a form of creeping resistance: weak resistance is augmented by stacking multiple resistance mechanisms (Duke et al., 2021).

1.2 Glyphosate resistance mechanisms

Glyphosate inhibits 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS) (Boocock and Coggins, 1983; Schönbrunn et al., 2001), a chloroplast enzyme involved in essential amino-acids (Tryptophan, Phenylalanine and Tyrosine) biosynthesis in the shikimate pathway (Bentley, 1990). Application of glyphosate on plants results in chlorosis, stunting, reduction in apical dominance and finally death (Baylis, 2000).

The mechanisms of glyphosate resistance have been extensively studied in microbes, weeds, and genetically modified crops. In this section, I examine the literature, focusing on microbes' and weeds' glyphosate resistance mechanisms with identified genetic bases. First, to detail the molecular mechanisms of glyphosate resistance, I first will address the cases of bacteria, fungi and yeast. Then I will present the glyphosate resistance mechanisms in weeds which focuses on the evolution of glyphosate resistance.

1.2.1 Glyphosate resistance mechanisms in microbes

Glyphosate inhibits EPSPS enzyme in the shikimate pathway, also referred to as the chorismate pathway, shared among plants, bacteria, fungi, algae and yeast (Hertel et al., 2021; Patriarcheas et al., 2023; Rong-Mullins et al., 2017).The widespread and intensive use of glyphosate also led to exposed microorganisms evolving resistance through a variety of mechanisms (Hertel et al., 2021; Hove-Jensen et al., 2014; Patriarcheas et al., 2023; Pollegioni et al., 2011). Here, I discuss more specifically glyphosate resistance mechanisms in microbes for which the molecular mechanisms have been investigated.

1.2.1.1 Target-site resistance mechanisms

Glyphosate target-site mechanisms involve modifications to the target enzymes (Aro1, AroM, AroA or EPSPS depending on the species).

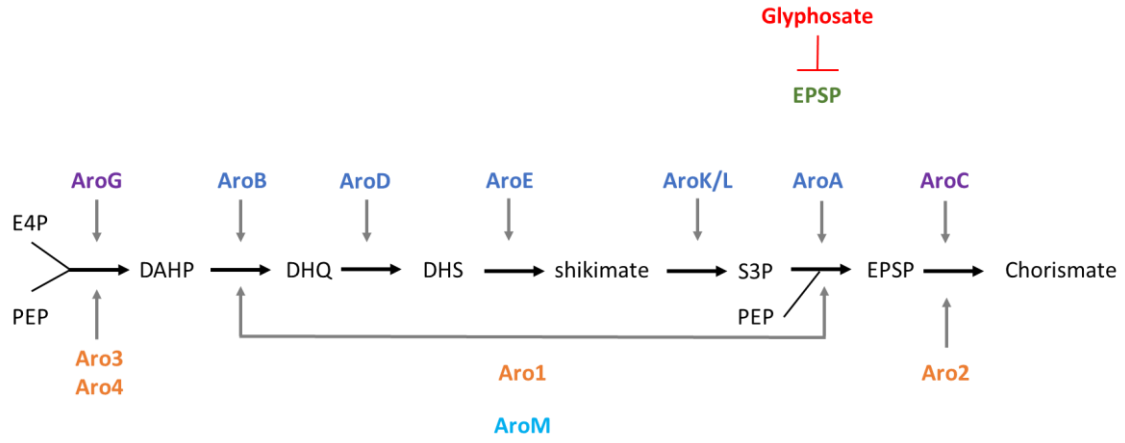

Figure 2 :Shikimate pathway (adapted from Patriarcheas et al.,2023). Yeast (orange), fungi (light blue), bacterial (dark blue and purple) and plant (green) genes coding for the enzymes involved are labelled. Compounds are in black. Glyphosate is structurally similar to, and competes with, PEP thus inhibiting EPSPS / AroA /Aro1 enzymes production of EPSPS. Full compounds names are: E4P= erythrose-4-phosphate, PEP= phosphoenolpyruvate, DAHP = 3-deoxy-D-arabino-heptulosonate-7-phosphate, DQH=3-dehydroquinate, DHS=3-dehydroshikimate, S3P=shikimate-3-phosphate, EPSP=5-enolpyruvylshikimate-3-phosphate*EPSPS, AroA, AroM* and *Aro1* genes are all related to the shikimate pathway, a seven-step biosynthetic cascade converting PEP and E4P into chorismate (Figure 2). Chorismate is crucial to the synthesis of the aromatic amino acids: tryptophan, phenylalanine and tyrosine. The *EPSPS* and the *AroA* genes both code for the 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme which catalyses one step of the shikimate pathway: the conversion of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) into 5-enolpyruvylshikimate-3-phosphate (EPSPS). These terms are sometimes used interchangeably but the *EPSPS* gene is primarily used in plants while the *AroA* gene is primarily used in bacteria and fungi. Other genes, *Aro1* and *AroM*, respectively found in yeast and fungi, code for a multifunctional enzyme complex that catalyses multiple steps of the shikimate pathway, including the steps performed by *EPSPS* and *AroA* genes (Campbell et al., 2004; Graham et al., 1993). Because glyphosate inhibits the activity of enzymes produced by these genes, any alteration to these enzymes endowing glyphosate resistance is classified as a target-site resistance (TSR) mechanism.

There are two main types of TSR mechanisms: gene overexpression or mutations of the gene encoding for the target enzyme. Overexpression can be due to either increased transcription or gene copy number variation (CNV). Usually, the terms overexpression and increased transcription are used interchangeably in the glyphosate resistant literature. To highlight the distinction between the molecular mechanisms at play, I have sub-categorised TSR glyphosate resistance as follows:

- Target-site mutation
- Target gene over expression due to either
    - Target gene copy number variation (CNV)
    - Target gene increased transcription

There have been accounts of Aro1 CNV endowing glyphosate resistance in the yeast *Saccharomyces cerevisiae* (Ravishankar et al., 2020a), the bacteria *Escherichia coli* (Wicke et al., 2019) and in a unicellular algae *Euglena gracilis* (Reinbothe et al., 1993). This mechanism of glyphosate resistance is largely present in weedy plants but has not yet been reported in fungi. Although the experimental evolutionary study by Ravishankar et al. (2020) reported Aro1 gene CNV is supporting glyphosate resistance in yeast, no mutations in the target gene were found. Conversely, the bacteria *E. coli* has been evolving glyphosate resistance through all three TSR mechanisms. Mutations rendering the AroA enzyme glyphosate insensitive include a single Arg102Ser substitution in isolation (Schwedt et al., 2023; Wicke et al., 2019), or in combination with up to five other amino acids changes (Schwedt et al., 2023). A mutation in serC, an AroA gene promoter, leading to AroA gene over expression and AroA gene duplication have also been documented (Wicke et al., 2019). Glyphosate resistance endowing mutations in AroA domain of the AroM complex of fungus *Acremonium sp* have also been evidenced through gene mining (Liu and Cao, 2018).

1.2.1.2 Non-Target-site resistance mechanisms

Glyphosate non-target-site resistance (NTSR) mechanisms includes all resistance mechanisms that do not directly impact the target enzymes. These mechanisms can result from a single process or a combination of processes, reducing glyphosate concentration at its site of action and thereby neutralizing its lethal effects (Tani et al., 2015).

1.2.1.2.1 Reduced uptake

GltP and GltT are dicarboxylate/amino acid cation symporters (Saier, 2000; Slotboom et al., 1999). GltT is a high-affinity glutamate/aspartate symporter encoded by *gltT* gene (Zaprasis et al., 2015) and GltP is a low-affinity glutamate transporter (Tolner et al., 1995). A recent study by Wicke et al. (2019) revealed that glyphosate resistance in *Bacillus subtilis* was acquired through a loss of function mutation in the *gltT* gene, sometimes accompanied by a second loss of function mutation in the *gltP* gene. Further genetic and metabolome analysis confirmed that GltT is a major glyphosate transporter and that GltP also mediates glyphosate transport in *B subtilis*. Contrary to the result found in *E. coli*, there was no mutation in AroA (Wicke et al., 2019). In summary, glyphosate resistance mechanisms seem to mainly involve NTSR mechanisms in *B subtilis,* while *E. coli* glyphosate resistance relied on three different TSR mechanisms.

Dip5 is from the amino acid-polyamine organocation (APC) superfamily of transport proteins (Saier,2000a) and has low substrate specificity. Inactivation for *dip5* gene increases glyphosate resistance suggesting that Dip5 is involved in glyphosate uptake in yeast (Decottignies et al., 2002). A QTL analysis revealed that Dip5 was involved in glyphosate resistance in *S. cerevisiae*, further functional validation confirmed that Dip5 deletion or mutation increased glyphosate resistance (Rong-Mullins et al., 2017). There was no evidence for Aro1 mutation supporting glyphosate resistance in yeast but evidence for an additional NTSR mechanisms was found and will be discussed in the section below.

1.2.1.2.2 Increased efflux

Pdr5 is an ABC efflux transporter whose inactivation increases glyphosate sensitivity, indicating its potential role in glyphosate export in yeast (Decottignies et al., 2002). In *S. cerevisiae* a QTL analysis

confirmed that Pdr5 is associated with resistant phenotypes (Rong-Mullins et al., 2017). Further functional validation demonstrated that Pdr5 inactivation incurred loss of glyphosate resistance (Ravishankar et al., 2020b). This body of evidence suggest that glyphosate efflux out of the cell through Pdr5 ABC transporters is a glyphosate resistance mechanism in yeast.

Screening for glyphosate resistance mechanisms in *Aspergillus oryzae* and in *E. coli* followed by functional validation uncovered the role of major facilitator secondary (MFS) proteins. MFS40 is an uncharacterised membrane protein with potential contribution to glyphosate tolerance in *Aspergillus oryzae* RIB40 isolate. Functional validation through expression of MFS40 in *E. coli* confirmed increased glyphosate tolerance. (Tao et al., 2017). The *yhhs* gene encoding for an MFS protein involved in drug efflux has been identified as a candidate in *E. coli*. Its overexpression in *E. coli* and *Pseudomonas* confers high levels of glyphosate resistance by reducing accumulation of glyphosate in the cells (Staub et al., 2012).

1.2.1.2.3 Detoxification

The metabolic detoxification pathways for xenobiotics—chemical compounds foreign to an organism's biochemical system—are of particular interest in understanding multidrug resistance. These pathways, which enable organisms to survive exposure to toxic substances, are also significant in an environmental context for assessing the feasibility of bioremediation in contaminated areas, as well as in agriculture for studying the evolution of pesticide resistance. Given its toxicity to microorganisms, glyphosate can be classified as a xenobiotic. The detoxification of xenobiotics typically involves a three-phase process, wherein lipid-soluble toxins are converted into water-soluble, neutralised metabolites (Dubey et al., 2014):

Phase I - Functionalization: In this phase, xenobiotics undergo modification to increase their polarity, thereby facilitating their preparation for subsequent processing in Phase II. This step is often mediated by monooxygenases, such as Cytochrome P450 enzymes.

Phase II - Conjugation: During this phase, the modified xenobiotic is conjugated to form a water-soluble compound, promoting its excretion in Phase III. The glutathione S-transferases are among the most extensively studied enzymes involved in this phase.

Phase III - Transport and Excretion: In the final phase, the conjugated compound is eliminated from the organism through the action of membrane transporters, such as ATP-binding cassette (ABC) transporters.

Exposure of the fungus *Aspergillus nidulans* to a no-observed-adverse-effect level (NOAEL) dose of a glyphosate-based herbicide resulted in differential expression of 1,816 genes. Subsequent analysis indicated that several of these differentially expressed genes are associated with functional groups involved in xenobiotic detoxification, including cytochrome P450 enzymes responsible for oxidoreduction reactions (Phase I), glutathione S-transferases (GST) catalysing the conjugation of reduced glutathione (Phase II), and ATP-binding cassette (ABC) transporters facilitating the export of conjugated xenobiotics (Phase III) (Mesnage et al., 2020) . These findings suggest that changes in the expression of these genes may represent *A. nidulans*' response to exposure to low concentrations of a herbicide containing glyphosate and surfactants. However, further validation is necessary to confirm their specific roles in glyphosate detoxification.

Detoxification through covalent modification is a common mechanism in Streptomyces species, and glyphosate undergoes a similar process via N-acetylation: the N-acetylated form of glyphosate does not inhibit the enzyme EPSPS in *Bacillus licheniformis* (Castle et al., 2004). This N-acetylation of glyphosate is catalysed by glyphosate acetyltransferase (GAT) enzymes (Castle et al., 2004) and has also been observed in Achromobacter species (Shushkova et al., 2016) Additionally, covalent modification of glyphosate can occur through ATP-dependent phosphorylation, a reaction catalysed by the enzyme Hph in *E. coli* and by GlpA in *Pseudomonas pseudomallei* (Penaloza-Vazquez et al., 1995).

1.2.1.2.4 Enhanced metabolic degradation

The compilation of bacteria and fungi capable of degrading glyphosate, along with the associated degradation pathways, has been comprehensively summarized by Feng et al. (2020) and Singh et al. (2020). The majority of research in this domain adopts a bioremediation perspective, focusing on leveraging microorganisms' capacity to degrade glyphosate for the rehabilitation of contaminated soil, with relatively fewer studies examining the genetic basis of glyphosate degradation. Consequently, most investigations have concentrated on identifying the specific degradation pathways involved, occasionally assessing environmental conditions that facilitate glyphosate biodegradation. While this is an exceptionally compelling field of study, it lies beyond the scope of this thesis. Instead, this review of the existing literature aims to provide a concise overview of the various pathways conferring glyphosate resistance in microorganisms, and to compare them to pathways uncovered in weeds and to highlight key studies that have explored the genetic mechanisms underlying glyphosate degradation.

Glyphosate metabolic degradation pathways in microorganisms have been identified and categorized into two main types: C-N bond cleavage and C-P bond cleavage (see Figure 3) and has been extensively reviewed (Feng et al., 2020; Hertel et al., 2021; Hove-Jensen et al., 2014; Patriarcheas et al., 2023; Singh et al., 2020; Sviridov et al., 2015). Glyphosate degradation occurs either through the action of an oxidase enzyme, which cleaves the carboxymethylene-nitrogen bond of glyphosate, yielding aminomethylphosphonic acid (AMPA) and glyoxylate, or via a C–P lyase enzyme, which directly cleaves the carbon-phosphorus bond to produce sarcosine. Additionally, C–P lyase enzymes are involved in the further degradation of AMPA through cleavage of its C–P bond (Singh et al., 2020). Glyphosate degradation through the C-P lyase pathway is induced by phosphorus deficiency and therefore rarely occurs in natural environments. Consequently, glyphosate is predominantly metabolised into AMPA which is subsequently released into the environment. However, certain bacterial strains possess the capacity to utilize AMPA as a source of phosphate, despite lacking the ability to directly degrade glyphosate itself (Sviridov et al., 2015).In some cases there was evidence for both glyphosate oxidoreductase (GOX) and C-P lyase pathway operating in parallel (Firdous et al., 2020; Zhao et al., 2015).

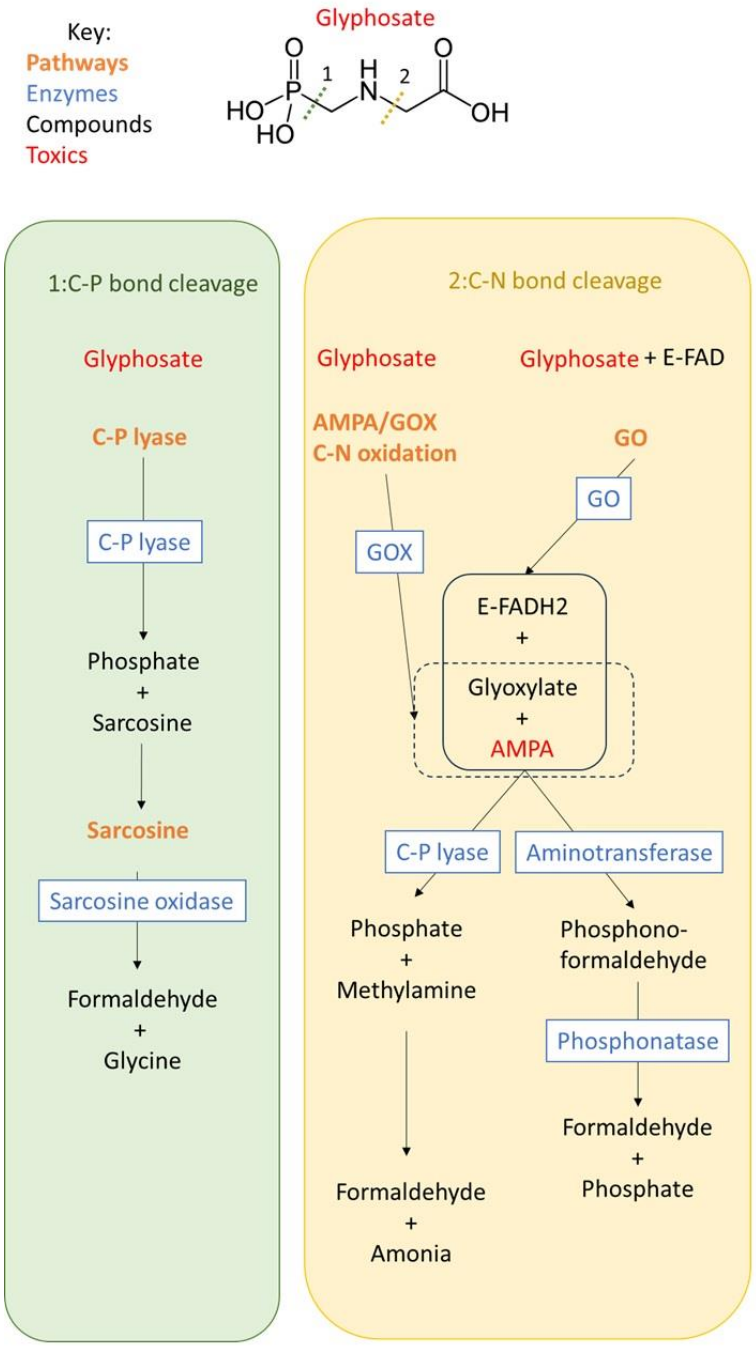

Figure 3: Glyphosate degradation pathways, adapted from Pollegioni et al, 2011 and Feng et al, 2020. 1: The C-P bond cleavage of glyphosate (also referred as the C-P lyase pathway) leads to the formation of sarcosine which can be further degraded through the sarcosine pathway. 2: The C-N bond cleavage of glyphosate can be catalysed either by the glyphosate oxidoreductase (GOX) or glycine oxidase (GO) enzymes (also referred to as oxidase dependent pathways). This leads to the formation of aminomethylphosphonic acid (AMPA) which is a toxic compound that will either be extruded in the environment or further metabolised by C-P lyases and Aminotransferases enzymes.

In bacteria, the C-P lyase pathway has been extensively studied and reviewed (Hove-Jensen et al., 2014; Stosiek et al., 2020). the Pho regulon constitutes a regulatory network that detects environmental phosphate levels and modulates gene expression to maintain phosphate homeostasis, facilitate phosphate uptake, and enable the utilization of alternative phosphorus sources during conditions of phosphate limitation (Santos-Beneit, 2015). A critical component of this network is the

phn operon, which is responsible for the degradation and utilization of phosphonates, such as glyphosate, as alternative sources of phosphate (Singh et al., 2020; Stosiek et al., 2020). In *E. coli*, the phn operon consists of a cluster of 14 genes that encode proteins essential for the uptake and cleavage of phosphonates, including key elements of a C-P lyase complex, which exhibits a high affinity for glyphosate. The three genes *phnCDE* encode an ABC transporter involved in phosphonate uptake (Chen et al., 1990; Hove-Jensen et al., 2011; Metcalf and Wanner, 1993), while the seven genes *phnG–I* and phnK–M encode the components of the C-P lyase complex, with *phnJ* functioning as a catalyst (Hove-Jensen et al., 2010; Metcalf and Wanner, 1993). Additionally, *phnF* encodes a repressor protein, and the three genes *phnNOP* are involved in accessory and regulatory functions within the C-P lyase pathway (Hove-Jensen et al., 2010; Metcalf and Wanner, 1993). Notably, there is evidence for the existence of another C-P lyase complex that coexists with the glyphosate-degrading C-P lyase and is involved in the degradation of AMPA (aminomethylphosphonic acid), a breakdown product of glyphosate (Hove-Jensen et al., 2010; Metcalf and Wanner, 1993).

A few studies screening for bacteria's ability to degrade glyphosate have explored the underlying genetic basis. In a study that screened soil isolates of *Pseudomonas pseudomallei* for their glyphosate-degrading capabilities, two candidate genes, *glpA* and *glpB*, were identified. Functional validation through expression in *E. coli* demonstrated their roles in glyphosate degradation: glpA contributes to glyphosate tolerance, while glpB facilitates the modification of glyphosate into a C–P lyase substrate. Additionally, the presence of hygromycin phosphotransferase (*hph*) genes from both *P. pseudomallei* and *E. coli* was found to confer glyphosate tolerance (Penaloza-Vazquez et al., 1995). Similarly, in *Pseudomonas* sp. isolated from contaminated soil, *glpA* was also identified as a candidate gene involved in glyphosate degradation, alongside *thiO*, *soxB*, and *argA*. In this strain, glyphosate degradation appears to be primarily mediated by the glycine oxidase (GO) enzyme (Zhang et al., 2024). Furthermore, in *Bacillus subtilis*, site-saturation mutagenesis of the GO enzyme led to the development of a variants with enhanced glyphosate-degrading capabilities. Specifically, a triple mutation (G51S/A54R/H244A) resulted in a 210-fold increase in catalytic efficiency compared to the wild-type enzyme (Pedotti et al., 2009). An RNA-seq study on glyphosate-tolerant fungi *Fusarium verticillioides* identified several differentially expressed genes (DEGs). Among these DEGs, some were found to encode an oxidoreductase, an ATPase, and a hydrolase. Notably, a gene related to glyphosate degradation, identified as part of the 3-isopropylmalate dehydratase of the aconitase superfamily (fv04), was further validated through heterologous expression in *E. coli*, demonstrating its role in glyphosate degradation (Guo et al., 2021).

Interestingly, an aldoketoreductase (AKR) ortholog, *igrA*, from the Pseudomonas strain PG2982 has been shown to confer glyphosate resistance (Fitzgibbon and Braymer, 1988, 1990). In this strain, the presence of a C-P lyase-mediated glyphosate degradation pathway was also demonstrated (Shinabarger and Braymer, 1986). This NADPH+/H+-dependent AKR is also found in plants, where it facilitates the conversion of glyphosate into AMPA (aminomethylphosphonic acid) and glyoxylate (Pan et al., 2019; Vemanna et al., 2017).To the best of my knowledge, this is the only evidence of a glyphosate degradation mechanism shared between plants and microorganisms.

1.2.2 Glyphosate resistance mechanisms in weeds

To better understand and manage glyphosate resistance in weeds the scientific community set out understand glyphosate resistance mechanisms. This extensive body of work has been summarised in a table presented in Appendix 1 collecting glyphosate resistance mechanisms found in weed species along with the methods employed. Here I review the different known mechanisms separately, however, they can co-exist in individuals and populations.

1.2.2.1 TSR

As previously, glyphosate target-site mechanisms involve any alterations to the target enzyme: EPSPS. Known mechanisms in weeds are *EPSPS* mutations, *EPSPS* overexpression and EPSPS increased enzyme activity.

1.2.2.1.1 *EPSPS* mutations

EPSPS mutation are changes to the *EPSPS* genes that alter EPSPS' affinity for glyphosate, rendering it less sensitive. Up to 13 mutations have been described in weeds (Collavo & Sattin, 2012; De Carvalho et al., 2012; Deng et al., 2022; Galeano et al., 2016; González-Torralva et al., 2012; Han et al., 2016; Jasieniuk et al., 2008; Li et al., 2018; Ng et al., 2003; Perotti et al., 2019; Takano et al., 2020) to date, with 8 single mutations, four double mutations and one triple mutation (see Appendix 1 for details and full reference list).

1.2.2.1.2 *EPSPS* overexpression

Gene expression can be influenced by regulatory regions (promoter, enhancer, silencer and insulator in eukaryotes), transcription factors, epigenetic regulation, post transcriptional processes, post translational processes (Singh et al., 2018) and gene copy number variation. The over expression of the *EPSPS* gene resulting in the increased abundance of the EPSPS enzyme induces glyphosate resistance since not all EPSPS enzymes will be inhibited by glyphosate. Two main mechanisms of *EPSPS* overexpression have been reported to date: increased transcription and copy number variation.

1.2.2.1.2.1 *EPSPS* increased transcription

Increased transcription is the over production of mRNA resulting in the increased production of the target enzyme. It can be constitutive or stress-induced. Evidence for *EPSPS* increased transcription has been confirmed in three weed species via RT PCR (see Appendix 1). The absence of known *EPSPS* mutations was confirmed (Mei et al., 2018; Yanniccari et al., 2017; Zhang et al., 2015). In the case of *Eleusine indica*, *EPSPS* copy number was approximately 5-fold higher in glyphosate resistant biotypes and the *EPSPS* gene expression was upregulated in a dose-dependent manner, contrary to that of glyphosate sensitive biotypes (Zhang et al., 2015). Constitutive increased transcription of the *EPSPS* gene in glyphosate resistant biotypes reaching up to 13-fold and 15-fold higher than in glyphosate sensitive biotypes has been evidenced in *Conyza canadensis* (Mei et al., 2018) and *Lolium perenne* (Yanniccari et al., 2017) respectively. The genetic basis of upregulation of *EPSPS* expression has not been fully resolved.

1.2.2.1.2.2 *EPSPS* copy number variation

Gene amplification or copy number variation (CNV) is another target-site mechanism of herbicide resistance whereby a genome contains multiple copies of a gene for a herbicide target, leading to increased production of the target enzyme to counteract the inhibitory effects of the herbicide (Bass and Field, 2011). CNV is a type of genomic structural variation that impacts expression without changes in the nucleotide sequence of the gene. *EPSPS* gene CNV is a glyphosate resistance mechanism that was first reported in *Amaranthus palmeri (Gaines et al., 2010)*. Convergent evolution of *EPSPS* CNV has occurred in up to 13 weed species (see Appendix 1). The increase in copies of the *EPSPS* gene can be located either in the genome or in extrachromosomal DNA (Patterson et al., 2018).

The first report of the genomic location with increased copies of *EPSPS* gene was found in glyphosate resistant *Kochia scoparia* with tandem amplification of a chromosomal segment leading to the presence of 10 copies of the *EPSPS* gene (Jugulam et al., 2014). *EPSPS* Cassette amplification was then described in *Amaranthus palmeri* glyphosate resistant biotype. The 297kb sequence includes the *EPSPS* gene, repetitive elements and putative helitron sequences (Molin et al., 2017). Authors suggest that an adaptative structural genomic mechanism drives the amplification and distribution of the *EPSPS* gene in the genome. Finally, subtelomeric rearrangements has been discovered as a mechanism increasing *EPSPS* gene copy numbers in the genome of glyphosate resistant *Eleusine indica* (Zhang, 2023). Chromosome telomeres and subtelomeres contain highly repetitive regions known to be prone to generating genomic structural variations. Authors found a unique rearrangement of the *EPSPS* gene that was inserted in one or more subtelomeric regions leading to an average 25 *EPSPS* copies in glyphosate resistant *E. indica*. They suggest that these duplication event occurred through unequal crossing over of subtelomere on chromosome three and potentially other chromosomes, but additional work is still required to confirm this hypothesis.

Recent work has uncovered other mechanisms leading to *EPSPS* CNV that is not located on chromosomes. Extrachromosomal circular DNA (EccDNA) is a type of double stranded circular DNA derived from chromosome but located outside chromosomes and was first discovered in 1965 (Hotta and Bassel, 1965). Since, efforts have been made to understand their biogenesis and function (Zuo et al., 2022). EccDNA containing copies of the *EPSPS* gene were found to confer glyphosate resistance in *Amaranthus palmeri* (Koo et al., 2018) and was transferred to *Amaranthus spinosus* by pollen flow during a hybridisation event *in natura*(Koo et al., 2023). EccDNA mediating *EPSPS* CNV was later found in *Lolium perenne ssp multiflorum (Koo et al., 2023)*. Future comparative analysis studies should shed light on origin and conservation of this particular EccDNA driving glyphosate resistance (Koo et al., 2023).

1.2.2.1.3 EPSPS increased enzyme activity

Glyphosate resistance was investigated in *Digitaria insularis* (Galeano et al., 2016). Two substitutions were present in the GR *EPSPS* gene: Pro15Thr and Tyr43Cys with the latter being essential to achieve glyphosate resistance. Comparison glyphosate resistant and glyphosate sensitive biotypes showed that constitutive *EPSPS* expression was on average slightly lower and constitutive EPSPS enzyme activity higher. There was no link between the EPSPS mutations and the gene expression or enzyme activity. Authors concluded that *D. insularis* glyphosate resistance mechanisms involved increased EPSPS enzymatic activity that is not directly related to gene expression nor potential NTSR mechanisms.

1.2.2.2 Non-Target-site mechanisms

Glyphosate non-target-site resistance (NTSR) encompasses all resistance mechanisms that do not directly affect the target enzyme EPSPS. These mechanisms may arise from a single process or a combination of processes, leading to a reduction in the concentration of glyphosate at its site of action and thereby preventing its lethal effects (Tani et al., 2015).

1.2.2.2.1 Reduced uptake

Reduced foliar uptake of glyphosate, due to changes in leaf morphology (shape, cuticular composition) can reduce the amount of glyphosate absorbed by the plant (Baek et al., 2021). This phenomenon has been detected in four different weed species using ^{14}C-glyphosate, showing differences in glyphosate uptake between susceptible and resistant biotypes (Michitte et al., 2007; Mora et al., 2019; Palma-Bautista et al., 2021; Vila-Aiub et al., 2011) The underlying molecular basis of reduced glyphosate uptake remains unknown to date.

1.2.2.2.2 Reduced translocation

Reduced translocation of glyphosate results in reduced transport of glyphosate to the meristems. (Baek et al., 2021). It can be achieved through glyphosate sequestration in vacuoles or transport outside of the cytoplasm.

1.2.2.2.2.1 Sequestration in vacuole

Vacuoles play a fundamental role in detoxification and maintain cell homeostasis. In glyphosate resistance, vacuolar sequestration acts to remove the glyphosate molecule from the cellular cytoplasm and stops additional translocation through the plant. It has been detected in six weed species (see Appendix 1) and is associated with a relatively low level of resistance (Baek et al., 2021). In some studies, the underlying molecular basis of vacuolar sequestration was not investigated and was solely evidenced by tracking ^{14}C-glyphosate (Michitte et al., 2007; Mora et al., 2019; Palma-Bautista et al., 2023; Vila-Aiub et al., 2011). Other studies demonstrated that ABC transporters are involved in this glyphosate resistance mechanism using transcriptome sequencing (Peng et al., 2010) or RT-PCR (Gerakari et al., 2022; Moretti et al., 2017; Peng et al., 2010; Schneider et al., 2021; Tani et al., 2015).

1.2.2.2.2.2 Transport outside of cytoplasm

Decreasing glyphosate cytoplasmic concentrations via transport outside of the cytoplasm was suspected to confer glyphosate resistance (Shaner, 2009). It has been observed in two weed species (see Appendix 1). The molecular basis for this mechanism has been evidenced in *Echinochloa colona* using RNA sequencing and RT-PCR uncovering the *EcABCC8* gene. The expression of this ABC transporter is upregulated in glyphosate resistant biotypes and endows glyphosate resistance when expressed in transformed rice. EcABCC8 is a plasma membrane–localized transporter extruding cytoplasmic glyphosate to the apoplast (Pan et al., 2021).

1.2.2.2.3 Glyphosate detoxification

In plants, metabolic herbicide resistance is often due to modification of herbicides by cytochrome p450s (CP450), glycosyltransferases, and glutathione S-transferase enzymes (GST), but functional validation for these enzymes has not yet been reported in cases of glyphosate resistance (Baek et al., 2021). Genomic regions enriched for genes in these families and associated with GR have been found in *Ipomoea purpurea* (Van Etten et al., 2020) but further validation is required for confirmation of their implication in glyphosate metabolic degradation (Baek et al., 2021). Similar studies found evidence for implications of ABC transporters, catalase (CAT), Peroxidase and superoxide dismutase (SOD) being enriched in glyphosate resistant biotypes transcriptomes (Deng et al., 2022; Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019a, 2019b). Ascorbate peroxidase (APX), glutathione reductase and dehydroascorbate reductase enzymatic activities have also been linked to glyphosate resistant weeds (Harre et al., 2018). All of these mechanisms remain to be functionally validated.

1.2.2.2.4 Enhanced metabolic degradation

Glyphosate metabolic degradation in weeds has only recently been discovered in three species (see Appendix 1). An aldo-keto reductase (AKR) enzyme was found to metabolize glyphosate to AMPA and glyoxylate in *Echinochloa colona* (Pan et al., 2019). Glyphosate resistant biotypes exhibited higher AKR gene expression and enzymatic activity than their glyphosate sensitive counterparts. Functional validation of EcAKR4-1 was performed in both rice and *E. coli* verifying AKR as a cause for glyphosate resistance. AKR4C10 were later discovered to endow glyphosate resistance in *E. indica* (Deng et al., 2022) and *Lolium rigidum* (Zhou et al., 2023), in the latter species AKR1 was also involved.

1.2.3 Concluding remarks on glyphosate resistant mechanisms

Both TSR and NTSR glyphosate resistance mechanism are shared amongst weeds, bacteria, fungi and yeast (Table 1). Target-site mutation have been reported to sustain glyphosate resistance in all taxa except yeast, in which it was investigated but not evidenced (Ravishankar et al., 2020b). Target gene increased transcription have been evidenced in weeds and bacteria but not in yeast and has not yet been reported in fungi. Target gene copy number variation has been reported in weeds, bacteria and yeast but not reported to date in fungi. Increased activity of the target enzyme (unknown molecular mechanisms) has been reported in one weed species (Galeano et al., 2016) and not reported in any other taxa.

NTSR mechanisms have been more consistently reported across weeds, bacteria, fungi and yeast with the exception of reduced uptake in fungi and detoxification and metabolic degradation in yeast that are not reported in the literature. Focussing on the molecular mechanisms underlying NTSR glyphosate resistance, different enzymes are reported in weeds and microbes with one exception: aldoketoreductase mediated glyphosate degradation was evidenced in *Pseudomonas sp* (Fitzgibbon and Braymer, 1988, 1990) as well as weeds (Deng et al., 2022; Pan et al., 2019; Zhou et al., 2023).

Table 1: Genetic basis of glyphosate resistance mechanisms in plants and in microbes. White cells indicate mechanisms that have either not yet been reported or investigated as sustaining glyphosate resistance in a particular taxon.

Mechanisms	Planta	Bacteria	Fungi	Yeast
TSR				
Target gene mutations	*EPSPS* single, double, TIPS, TIPT and triple (TIAVPS). For further information, see Appendix 1	*aroA* mutations in *E. coli* single mutations: Arg102Ser (Schwedt et al., 2023; Wicke et al., 2019) Multiple mutations: 37 Arg102Ser + at least one more mutation (for details see Table 2 in Schwedt et al, 2023)	AroA domain from the AroM complex GR variant in *Acremonium sp* (Liu and Cao, 2018)	*aro1* gene: no mutations found (Ravishankar et al., 2020b)
Target gene increased transcription	Present in three species but underlying mechanisms remain unknown (Mei et al., 2018; Yanniccari et al., 2017; Zhang et al., 2015)	*serC* (*AroA* promoter) mutation in *E. coli* (Wicke et al., 2019)	AroM complex or AroA domain increased transcription not yet reported	*aro1* gene increase expression not found (Ravishankar et al., 2020b)
Target gene copy number variation (CNV)	*EPSPS* gene tandem duplication (Jugulam et al., 2014) *EPSPS* cassette (Molin et al., 2017) Subtelomeric CNV (Zhang, 2023) eccDNA (Koo et al., 2018)	*aroA* gene duplication in *E. coli* (Wicke et al., 2019)	AroM complex or *aroA* domain CNV not yet reported	*Aro1* duplication (Ravishankar et al., 2020b)
Increased target enzyme activity	Present in one species but unknown mechanism (Galeano et al., 2016)	AroA increased enzyme activity not yet reported	AroM complex or AroA increased enzyme activity not yet reported	Aro1 enzyme increased activity not yet reported
NTSR				
Reduced uptake	Present in four species but underlying mechanisms are still unknown (Michitte et al., 2007; Mora et al., 2019; Palma-Bautista et al., 2021; Vila-Aiub et al., 2011)	Glyphosate transporters in *Bacillus subtilis* (Schwedt et al., 2023; Wicke et al., 2019) GltT GltP	No reduced uptake mechanisms reported yet	Dip5 transporter (Decottignies et al., 2002; Ravishankar et al., 2020b; Rong-Mullins et al., 2017)

Mechanisms	Planta	Bacteria	Fungi	Yeast
Reduced translocation/efflux	ABC transporter mediated Vacuolar sequestration (Gerakari et al., 2022; Moretti et al., 2017; Peng et al., 2010; Schneider et al., 2021; Tani et al., 2015) Cytoplasm efflux :EcABC8 (Pan et al., 2021)	*yhhs* encoding for MFS protein involved in drug efflux (Staub et al., 2012)	MFS40 uncharacterised membrane protein (Tao et al., 2017)	Pdr5 transporter (Ravishankar et al., 2020b; Rong-Mullins et al., 2017)
Detoxification	Still lacking functional validation: Cytochrome P450 (CP450) (Deng et al., 2022; Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019b) Glycosyltransferase (GT) (Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019b) GST (Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019b) ABC transporters (Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019b) Catalase (CAT) (Piasecki et al., 2019a, 2019b) Peroxidase (POD)(Piasecki et al., 2019b) Superoxide dismutase (SOD)(Piasecki et al., 2019a, 2019b) Ascorbate peroxidase (APX)(Piasecki et al., 2019a) Glutathione reductase (GR) (Harre et al., 2018) Dehydroascorbate reductase (DHAR) (Harre et al., 2018)	Glyphosate acetyltransferase (GAT) in *Bacilus lichenformis* and *Achromobacter sp* (Castle et al., 2004; Shushkova et al., 2016) Hygromicine phosphotransferase (Hph) in *E. coli* (Penaloza-Vazquez et al., 1995) GlpA in *Pseudomonas pseudomallei* (Penaloza-Vazquez et al., 1995)	Functional groups involved in detoxification with differential gene expression, still requiring further functional validation (Mesnage et al., 2020): Cytochrome P450 (CP450) Glutathione S-transferase (GST) ATP dependent binding Cassette (ABC) transporters	No detoxification mechanisms reported yet

Mechanisms	Planta	Bacteria	Fungi	Yeast
Enhanced metabolic degradation	aldo-keto reductase present in 3 species. EcAKR4-1 functionally validated (Deng et al., 2022; Pan et al., 2019; Zhou et al., 2023)	C-P lyase pathway: phn operon in *E. coli* (Hove-Jensen et al., 2014) *glpB* in *Pseudomonas pseudomallei* (Penaloza-Vazquez et al., 1995) Glycine oxidase (GO) in *Pseudomonas sp* and *Bacilus subtilis* (Pedotti et al., 2009; Zhang et al., 2024) Other candidate genes in *Pseudomonas* sp (Zhang et al., 2024): *thiO* *soxB* *argA* Aldo-keto reductase: *IgrA* (AKR1) in *Pseudomonas sp* (Fitzgibbon and Braymer, 1988, 1990)	Differentially expressed genes in *F. verticillioides* (Guo et al, 2021) Fv04 (functionally validated) Oxidoreductase ATPase Hydrolase	No degradation mechanisms reported yet

1.3 Experimental evolution

Experimental evolution is a method that allows the observation of evolution in real time directly testing evolutionary theories experimentally rather than relying solely on inferences from fossil records and the observation of evolutionary endpoints (Buckling et al., 2009). This approach involves applying natural selection in controlled environments by exposing replicate populations of organisms to novel environments while some are kept in ancestral environment to serve as controls (Garland and Rose, 2009). Researchers can precisely manipulate specific factors in the novel environment—whether biotic, abiotic, or demographic—and, through direct comparison, assess the effects on adaptation at various stages of the selection experiment. While experimental evolution has been successfully applied to multicellular organisms, it is particularly effective with microorganisms, which are easier to work with (Buckling et al., 2009).

1.3.1 Microbial Experimental Evolution (MEE)

MEE leverages characteristics of microbes such as their small size and short generation time to impose selection on large populations across many generations under controlled condition (Buckling et al., 2009; McDonald, 2019), hereby effectively mitigating two primary experimental constraints: limited available time and space. Additionally, most model microorganisms have relatively simple and well-known genomes, allowing for genetic manipulations and investigation of the genetic basis of adaptation (Bell and Reboud, 1997; Buckling et al., 2009; Elena and Lenski, 2003). Lastly, the possibility of storing living cells allows the creation of a "living fossil record" (Lenski et al., 1991), to further test the effects of selection on the intermediate phenotypes and genotypes.

1.3.1.1 MEE as a tool to investigate fundamental evolutionary questions

Microbial experimental evolution (MEE) has provided empirical insights into several aspects of evolutionary theory. It has provided clear evidence of natural selection in microbial populations by observing adaptation dynamics and the role of beneficial mutations in driving evolutionary change (McDonald, 2019). MEE studies have demonstrated the partial predictability of evolution, as similar traits and mutations arise independently across populations under consistent selective pressures (Cooper et al., 2003; Herron and Doebeli, 2013; Lang et al., 2013; Wichman et al., 1999). Such studies also shed light on epistasis, revealing how mutation effects depend on existing genetic backgrounds, often leading to diminishing returns in already adapted populations (Kryazhimskiy et al., 2009; Wiser et al., 2013). Additionally, MEE explores the balance between adaptation and genetic drift, with small populations showing a greater influence of drift (Elena et al., 2007; Tenaillon et al., 1999; Wilke, 2004; Willi et al., 2006). The effect of genetic diversity on the evolutionary outcome has been investigated by using starting populations with large amounts of standing variation (Burke et al., 2014; Jerison et al., 2017; Kosheleva and Desai, 2018), a genetic clone (Kao and Sherlock, 2008; Rainey and Travisano, 1998) or artificial elevation of mutational rates using mutagens or disabling genes involved in DNA repair (Gray and Goddard, 2012; Koschwanez et al., 2013). Whole population sequencing at multiple time points has enabled tracking of individual mutations arising and segregating in evolving populations (Behringer et al., 2018; Good et al., 2017; Lang et al., 2013). These studies provided a direct view of mutation trajectories and interactions during adaptation and allowed to observe and explain how clonal interference slows adaptation (Good et al., 2017; Lang et al., 2013). Together, these findings from MEE bolster and refine classical evolutionary theories with real-time, molecular-level data (McDonald, 2019). Additionally, an MEE experiment accessing long

evolutionary timescales by leveraging *E. coli's* short generation time showed that populations still evolved to their laboratory conditions after over 61,500 generations (Good et al., 2017; Wiser et al., 2013) while theory would have predicted that optimal fitness would be reached in a few thousand generations.

Early experimental evolution studies primarily focused on adaptation to a specific ecological scenario: the occurrence of a single, abrupt environmental change followed by long-term adaptation to a new, stable environment (Collins, 2011; Collins et al., 2007). Subsequent research expanded this approach by investigating adaptation in more complex ecological contexts, such as fluctuating environments (Jessup et al., 2004). The rate of environmental change has been shown to influence both the evolutionary outcomes and the repeatability of evolution, shaping whether generalists or specialists evolve and driving adaptive divergence among replicate populations (Buckling et al., 2007; Condon et al., 2014). In the specific context of adaptation to directionally changing environments (e.g. where populations are exposed to a sequence of progressively deteriorating conditions), the rate of change affects both the risk of extinction (Bell and Collins, 2008; Chevin et al., 2010) and the nature of adaptive responses if the population survives. Studies have demonstrated that populations adapting to slower rates of environmental change achieve higher final fitness compared to those adapting to faster rates (Collins and De Meaux, 2009; Collins et al., 2007). Adaptation to slower rates of environmental change is associated with smaller incremental fitness gains, suggesting that different mutations are fixed compared to populations experiencing faster environmental change. Simulation studies (Collins et al., 2007; Kopp and Hermisson, 2007) suggest that slower rates of change reduce the likelihood of large fitness drops, leading to the fixation of mutations with smaller effects. Notably, while Kopp and Hermisson (2007) assumed a predetermined endpoint, (Collins et al. (2007) found that evolved populations exhibited different final phenotypes under varying rates of change, aligning with later empirical findings (Collins and De Meaux, 2009). Two studies investigating the effect of rate environmental change on the adaptation of yeast populations to different heavy metals demonstrated that the evolutionary pathways and the genomic adaptation were influenced by the rate of environmental change and the nature of the stressor (Gorter et al., 2016, 2017).

1.3.1.2 MEE bridging fundamental insights and applied strategies: the case of antibiotic resistance

Microbial experimental evolution (MEE) is a robust approach to investigate fundamental evolutionary questions, while also addressing the pressing applied issue of antibiotic resistance (McDonald, 2019). The evolution of antibiotic resistance is a clear example of rapid adaptation and poses a significant threat to global health (Jansen et al., 2013; Perron et al., 2008) , primarily because it often results in treatment failure (Jansen et al., 2013). While comparative genomics facilitates the identification of genetic mechanisms underlying resistance, MEE provides a systematic and controlled framework for dissecting the evolutionary trajectories that give rise to problematic resistant strains (Jansen et al., 2013). Gaining insight into these evolutionary pathways is critical for anticipating the emergence of resistance (Buckling et al., 2009) and developing sustainable management strategies (Jansen et al., 2013).

Most resistance management strategies relied on resistance mechanisms being associated with a fitness cost (Andersson and Hughes, 2010). MEE studies have quantified the fitness costs associated with antibiotic resistance mutations (Chevereau et al., 2015; Denamur et al., 2005; MacLean and Buckling, 2009), offering insights into the trade-offs that influence resistance dynamics. However, research indicates that the effects of these mutations and their associated costs are not always predictable (Kassen and Bataillon, 2006; MacLean and Buckling, 2009; Nang et al., 2018; Rozen et al.,

2007). Furthermore, initial fitness costs may be mitigated through secondary compensatory mutations (Levin et al., 2000). The application of antimicrobial compounds often creates complex spatiotemporal selection gradients, which are likely to influence the emergence and spread of antibiotic resistance (Levy and Marshall, 2004; O'Brien, 2002). Consequently, MEE has been used to study the effects of heterogeneous antibiotic environments. For example, *E. coli* populations exposed to antibiotic concentration gradients showed that adaptation to high concentrations resulted in high fitness across the gradient, whereas populations exposed to lower concentrations exhibited initial fitness disadvantages at higher concentrations but adapted more rapidly over time (Lagator et al., 2021).

The directional changes in antibiotic concentrations caused by the accumulation of antimicrobial compounds in clinical, veterinary, and agricultural settings further drive bacterial adaptation (Levy and Marshall, 2004). A notable MEE study examined the interplay between immigration rates and environmental changes in shaping the evolution of resistance, revealing that rapid environmental shifts combined with high immigration rates resulted in higher levels of resistance (Perron et al., 2008).

Just as antibiotic resistance threatens human health, herbicide resistance poses an increasing challenge to global food security. These instances of human-induced resistance evolution share significant parallels (Beckie et al., 2021). Leveraging MEE to enhance understanding of herbicide resistance and inform management strategies is feasible, particularly through the use of the model organism *Chlamydomonas reinhardtii*. However, while the use of a unicellular, asexually reproducing model organism to study the evolution of herbicide resistance provides valuable insights, the findings may not be directly transferable to higher plants. In unicellular systems, resistance mechanisms such as reduced uptake or enhanced exclusion are likely to play a more prominent role than in multicellular organisms. Furthermore, sexual reproduction in higher plants is expected to facilitate the dissemination of advantageous traits, such as herbicide resistance, under selection pressure. Nevertheless, employing *C. reinhardtii*, a model organism with many desirable characteristics to study evolution of herbicide resistance enables the investigation of common fundamental evolutionary dynamics underlying glyphosate resistance in this thesis. Generally, MEE studies offer a foundation for formulating hypotheses that can be subsequently tested in more complex and ecologically relevant systems.

1.3.2 Leveraging MEE to study the evolution of herbicide resistance

Although the model plant species *Arabidopsis thaliana* has been extensively used to study herbicide resistance (Jander et al., 2003; Roux and Reboud, 2005; Roux et al., 2004, 2005), such studies require substantial resources to maintain plants under controlled conditions over multiple generations. Unicellular algal model species present a more practical alternative for investigating the evolution of herbicide resistance. Notably, *C. reinhardtii*, a member of the Viridiplantae (green plants), shares a common ancestry and physiological traits with higher plants. Its cellular biochemistry being similar to that of higher plants, *C. reinhardtii* has already been employed as a model to explore herbicide resistance (Fedtke, 1991).

1.3.2.1 Chlamydomonas reinhardtii, a model species

1.3.2.1.1 *C. reinhardtii* characteristics

C. reinhardtii is a cosmopolitan unicellular flagellated eukaryote (Chlorophyceae) with a global natural distribution across both soil and freshwater ecosystems (Melero-Jiménez et al., 2021; Nestler et al., 2012a). With laboratory strains available from the *Chlamydomonas Resource Center* (https://www.chlamycollection.org/), *C. reinhardtii* serves as a model organism for a wide range of research fields, owing to a range of advantageous traits (Dupuis and Merchant, 2023). This species exhibits a short generation time, capable of achieving up to 10–12 generations per week under optimal conditions, and large populations can be cultured following well-documented protocols. It can be maintained as haploid vegetative cells and can also be cultured in liquid media or stored on agar. Depending on environmental conditions, *C. reinhardtii* exhibits heterotrophic and facultative autotrophic modes of nutrition and is capable of both sexual and asexual reproduction and these conditions are readily controllable in a laboratory setting (Harris, 2008a). Furthermore, *C. reinhardtii's* genetic tractability is enhanced by the ability to reveal loss-of-function phenotypes through mutagenesis, while its capacity for sexual reproduction enables the application of classical genetic approaches. Additionally, its genome has been fully sequenced (Merchant et al., 2007). These features collectively establish *C. reinhardtii* as a model species of choice for molecular and microbial experimental evolution studies.

1.3.2.1.2 *C. reinhardtii* genome

C. reinhardtii genome was first published by Merchant et al. (2007). Following iterative genome assembly progress and gene model refinement, several updated versions of this genome were made available on Phytozome, the Joint Genome Institute's (JGI) plant genomic portal (Blaby et al., 2014). This genome was assembled from sequencing a cell wall-less strain of mating type +(CC-503) and uncovered a non-compact genome (~110MB) with genes carrying on average 7 introns (of >350 bp) (Merchant et al., 2007), and relatively active transposable elements of class I and II (Gallaher et al., 2015; Kim et al., 2006; O'Donnell et al., 2020). While the nuclear genome is GC rich (~64%), the organelle genomes (34.5% in the chloroplast and 45.2% in the mitochondrial genomes) have a much lower GC content (Ness et al., 2012). There are on average 83 copies per cell of the highly repetitive ~250 kbp circular chroroplast genome and 130 copies per cell of the ~15.8 kbp linear mitochondrial genome (Gallaher et al., 2018).

Using the *C. reinhardtii* reference genomes version published on Phytozome as a reference, a series of studies employed mutation accumulation experiments to investigate spontaneous mutations in *C. reinhardtii*, shed light on mutations that arise in the absence of selection (Morgan et al., 2014; Ness et al., 2012, 2015). The initial study, conducted with a single strain, identified 14 spontaneous mutations over ~350 generations, estimating a total mutation rate of 3.23 $x10^{-10}$ mutations. $site^{-1}$. $generation^{-1}$ and the single base mutation rate of 2.08 $x10^{-10}$ $mutations.base^{-1}.generation^{-1}$ (Ness et al., 2012). Further research on multiple strains revealed significant variation in mutation rates among strains (Ness et al., 2015). Overall, deletions were more common than insertions, and fine-scale variation in mutation rates was observed (Ness et al., 2015). Authors developed an effective

predictive model for mutation rates based on genomic features, suggesting that spontaneous mutations are more likely to occur in the immediate vicinity of high GC regions, near specific trinucleotide sequences (such as CTC) and loci associated with higher transcription levels (Ness et al., 2015). Despite the high GC content of the nuclear genome (~64%), a paradoxical G/C to A/T mutation bias was detected, suggesting a prominent role for widespread biased gene conversion in the nuclear genome (Ness et al., 2012). Another study investigated the fitness effect of spontaneous mutations establishing that 5.6% of mutations were deleterious with an average effect of -4.07x10-2 μmax generation-1 (Morgan et al., 2014).

The latest version of assembly and annotation of this genome is version V5.6. At the time, the technology did not allow sequencing of entire chromosomes which inevitably lead to gaps in the assembly. Major issues were caused by the presence of repeats (identical sequences occuring in multiple genome locations) of a greater size than the sequenced reads (making it impossible to know the copy from which the reads originated) and regions such as high GC from wich it is difficult to obtain the sequence (Blaby et al., 2014). Despite over a decade of improvements to the assembly (Blaby et al,2014), the Phytozome-CC503_V5.6 *C. reinhardtii* genome is contained in 54 scaffolds (17 chromosomes and 37 minor scaffolds) assembled from 1,495 contigs (Merchant et al., 2007). and ~2Mb of sequence remains unplaced.

Strain variation, in addition to assembly quality, poses challenges for genomic analysis. Laboratory strains of *C. reinhardtii* were traditionally classified into three main lineages: the Sager, Cambridge, and Ebersold-Levine lineages, all descended from a field-isolated strain collected by Smith in 1945 (Harris, 2008b). However, recent genomic comparisons of 39 strains suggest that some strains have been misidentified, and a five-lineage model (lineages I to V) provides a more accurate classification framework (Gallaher et al., 2015).

In this thesis, I utilized the *C. reinhardtii* strain CC-1690, which belongs to the Sager lineage in the three-lineage model and Lineage III in the five-lineage model. Although CC-1690 and CC-503 (from the Ebersold-Levine lineage) share a relatively recent common ancestry and most of their genomes are identical by descent, significant polymorphism exists between them (Flowers et al., 2015; Gallaher et al., 2015).

Comparative analyses revealed approximately 61,480 single nucleotide polymorphisms (SNPs) distributed unevenly across the genome (on chromosomes 2, 4, 8, 9, 10, 11, 15, and 16). Additionally, entire chromosomal segments were found to be non-identical by descent, and certain regions of the CC-1690 genome failed to map onto the reference genome (Flowers et al., 2015). De novo reassembly of these unmapped regions led to the identification of 12 potential novel genes (Flowers et al., 2015). Large-scale duplications, characterized by abnormally high genomic coverage compared to the average, emerged as the most prominent class of mutations (Flowers et al., 2015). These copy number gains were localized to specific regions of chromosome 13, including loci 4,141,500–4,227,500, 4,349,500–4,403,500, and 4,487,000–4,537,000 (Flowers et al., 2015). Furthermore, large deletions were frequently associated with transposon positions, suggesting that transposon activity contributes significantly to structural variation within the genome (Gallaher et al., 2015).

To address limitations arising from both strain variation and assembly quality, researchers have opted to perform de novo genome assembly for the specific strains under study using long-read sequencing technologies (Payne et al., 2023). In 2020, a highly contiguous nuclear genome assembly of the

laboratory strain CC-1690 was published using nanopore sequencing (O'Donnell et al., 2020). This assembly spans 111 MB, consists of 21 contigs, and includes five additional complete Benchmarking Universal Single-Copy Orthologs (BUSCOs) compared to the latest CC-503 reference genome available on Phytozome (O'Donnell et al., 2020).

1.3.2.1.3 *C. reinhardtii* and MEE studies

Due to its microbial-like experimental advantages, *C. reinhardtii* serves as an excellent model for investigating fundamental biological questions and has been extensively used in MEE studies since the 1990s.

Several studies have investigated the evolutionary dynamics of *C. reinhardtii* populations under various ecological conditions. Research has explored the impact of environmental heterogeneity on genetic variation (Bell, 1997; Bell and Reboud, 1997) and evolution of generalist and specialist strategies (Kassen and Bell, 1998; Reboud and Bell, 1997). Other studies have assessed the effects of population size on the repeatability of adaptation (Lachapelle et al., 2015a), the role of sexual reproduction in evolutionary rescue under deteriorating environments (Lachapelle and Bell, 2012) and the influence of selection history on extinction risk during severe environmental changes (Lachapelle et al., 2017). *C. reinhardtii* mutation rate has been estimated (Ness et al., 2012) and the process of spontaneous mutation accumulation has been described (Böndel et al., 2019; Morgan et al., 2014b; Ness et al., 2015).

While much of this research has focused on the effects of stable environmental changes, recent studies have begun to examine the consequences of variable environmental conditions. For instance, investigation of the impact of environmental rates of change on adaptive outcomes and dynamics, including fitness costs and types of mutations (Collins and De Meaux, 2009). Or studying the interplay between mode of reproduction and extinction dynamics in response to varying rates of environmental deterioration (Petkovic and Colegrave, 2023).

1.3.2.2 Previous studies on herbicide resistance in *C. reinhardtii*

An extreme case of adaptation to environmental change is the adaptation to herbicide exposure. As a common primary producer, *C. reinhardtii*'s response to herbicide serves as a relevant model for understanding their broader impact on non-target components of the agricultural ecosystem (Melero-Jiménez et al., 2021; Nestler et al., 2012a). This species' shared cellular biochemistry with higher plants and susceptibility to herbicide (Reboud, 2002) further establish *C. reinhardtii* as a valuable model for studying the evolution of herbicide resistance. Moreover, a study employing genetically engineered herbicide-resistant *C. reinhardtii* proposed that this approach could be applied for crop protection in algal production systems (Bruggeman et al., 2014). However, if such techniques are adopted in algal production, there is a likelihood that herbicide resistance will evolve in undesirable organisms, mirroring the patterns observed in modern agricultural cropping systems. This underscores the importance of studying the evolution of herbicide resistance in *C. reinhardtii*.

Three main experimental designs have been employed to study the evolution of herbicide resistance in *C. reinhardtii*, each imposing distinct constraints on the system. Ratchet protocols, where populations are transferred to fresh media upon achieving sufficient growth (Melero-Jiménez et al., 2021; Reboud et al., 2007; Vogwill et al., 2012), minimize bottleneck effects. However, if

environmental conditions change during the transfer to fresh media, these protocols can also slow the rate at which change in selective pressure is applied. Source-sink scenarios (Kawecki and Holt, 2002; Lagator et al., 2012, 2013, 2014) involve regular transfers into fresh media but mitigate bottleneck effects by allowing immigration from a source population when required. Finally, continuous flow cultures in mesostats eliminate bottlenecks effects altogether, while maintaining precise control over nutrient concentrations throughout the experiment (Hansson et al., 2022).

Herbicide resistance endowing mutations were originally evidenced in Chlamydomonas by a number of early studies (Erickson et al., 1984, 1989; Fedtke, 1991; Galloway and Mets, 1984; Hartnett et al., 1987; James et al., 1993; Randolph-Anderson et al., 1998). In light of this, the use of *C. reinhardtii* as a model organism for studying the evolution of herbicide resistance was initiated by a study that established dose-response curves for 29 herbicides (Reboud, 2002). This was subsequently followed by the development of a protocol to evolve resistance to atrazine under controlled conditions, demonstrating the suitability of *C. reinhardtii* for experimental evolution studies of herbicide resistance (Reboud et al., 2007). These pioneering studies opened the door for further research, focussed on understanding the evolution of herbicide resistance in *C. reinhardtii*, the effects of different selective environments, and the implications for resistance management strategies in an agricultural context. The aim was to take an evolutionary biology approach by studying the effect of management practices (herbicide sequential application, cycling, mixture and dose) and their underlying ecological and evolutionary theories (Lagator et al., 2012, 2013, 2014; Vogwill et al., 2012).

The relationships between herbicide resistance, fitness in the ancestral environment, and the effects of different herbicide treatments have been explored in various studies (Lagator et al., 2012, 2013, 2014; Vogwill et al., 2012). A key finding from this body of work is that *C. reinhardtii* populations exposed to herbicides such as atrazine, glyphosate, and carbetamide consistently evolve resistance (Lagator et al., 2012, 2013, 2014; Vogwill et al., 2012). Research has also investigated the efficacy of herbicide mixtures in delaying or preventing the evolution of resistance, finding that herbicide mixtures at high doses can slow resistance evolution, whereas mixtures at low doses may accelerate resistance evolution and promote cross-resistance (Lagator et al., 2013). Cross-resistance has been frequently observed in *C. reinhardtii* populations that evolved resistance to herbicides under herbicide cycling or mixing strategies (Lagator et al., 2012, 2013) while the impacts of selection history on the dynamics of adaptation under sequential herbicide exposure have also been examined, with pre-exposure to one herbicide often facilitating rapid adaptation to another (Lagator et al., 2014).

Where herbicide resistance has been experimentally evolved, the impact on this on subsequent fitness in either the presence or absence of the herbicide can be investigated, allowing an empirical assessment of potential resistance-associated fitness costs. Resistance-associated fitness costs have been described for *C. reinhardtii*, and shown to be affected by the frequency of cycling between herbicide actives (Lagator et al., 2012). More broadly, results have demonstrated that the relationship between herbicide resistance and fitness costs is complex and context-dependent (Lagator et al., 2012; Vogwill et al., 2012). Overall, the evolutionary dynamics of herbicide resistance were shown to vary across herbicides and experimental conditions, indicating that no universal resistance management strategy can be recommended. A notable conclusion is that management strategies tested often favoured the evolution of generalist phenotypes, with their effects on fitness costs remaining unpredictable.

Recent studies have investigated the effects of glyphosate selective doses on the evolution of resistance in *C. reinhardtii*. To allow between studies comparison I will refer to glyphosate selective doses in relation to the minimum inhibitory concentration (MIC) which is the lowest dose required to totally inhibit growth. (Hansson et al., 2024) examined resistance evolution under continuous exposure to glyphosate at either a lethal dose (1 MIC) or a sublethal dose (0.5 MIC), representing a rapid rate of glyphosate dose escalation. Under 1 MIC, a marked population decline was observed at the onset of selection, whereas populations exposed to 0.5 MIC did not exhibit this initial decline. Rapid evolution occurred under 1 MIC, with recovery observed within 19–22 days post-exposure. Growth rate assays conducted across a glyphosate gradient (from 0 up to 1.5 MIC) indicated delayed resistance evolution in populations selected under 0.5 MIC. Notably, no significant shift in MIC or fitness costs associated with evolved glyphosate resistance was detected under either selection regime (Hansson et al., 2024). In contrast, (Melero-Jiménez et al., 2021) investigated the effects of incremental glyphosate dose increases using a ratchet protocol, (i.e. starting at 0.1 MIC and increasing the dose once adaptation occurred), representing a slow pace of glyphosate dose escalation. Similarly to Hansson et al. (2024), their study demonstrated the evolution of glyphosate resistance. However, they observed a shift in levels of resistance up to 1.8 MIC in selected populations, alongside evidence of fitness costs associated with resistance. These costs were characterized by reduced growth rates in the ancestral environment and impaired photosynthetic performance relative to control populations (Melero-Jiménez et al., 2021). The apparent discrepancies between these findings may be explained by differences in the pace of environmental change, which can profoundly influence evolutionary outcomes (Collins and De Meaux, 2009; Collins et al., 2007; Perron et al., 2006, 2008).

1.3.2.3 Harnessing MEE and Next-Generation Sequencing with *C. reinhardtii* to uncover the genetic mechanisms of herbicide resistance

Experimental evolution studies model evolutionary processes with convenient experimental systems, traditionally characterising evolution at the phenotypic level. However, the advent of next generation sequencing (NGS), and particularly the drop in sequencing costs, enabled researchers to link observed phenotypic responses to underlying genetic changes (Brockhurst et al., 2011; Schlötterer et al., 2015). Such studies are now being used to precisely estimate mutation rates, identify genetic targets and the dynamics of natural selection, explore the relationship between genetic and phenotypic changes, and test long-standing evolutionary hypotheses (Brockhurst et al., 2011).

In the context of herbicide resistance, the International Weed Genomics Consortium has emphasized the importance of understanding the molecular basis of herbicide resistance evolution as a key research priority (Ravet et al., 2018). However, most of the work presented here focused on observations of the evolved herbicide resistant phenotype and insight would be gained from understanding the genetic mechanisms underpinning the evolution of herbicide resistance (Lagator, 2012a). Combining experimental evolution with NGS presents a promising approach to address this gap.

Although no studies have yet applied microbial experimental evolution (MEE) in combination with NGS to investigate herbicide resistance evolution, existing research illustrates the potential of this approach. For example, (Kronholm et al., 2017) combined MEE with genomics and methylomics to

study the genetic and epigenetic contributions to adaptation to stressful environments (high salt, high CO_2, and low phosphate) in *C. reinhardtii*. Their findings revealed that epigenetic mutations occur at a faster rate than genetic mutations, can be inherited, and contribute to adaptation in stressful environments, with the extent of their contribution being environment-dependent (Kronholm et al., 2017).

In addition to epigenetics, genetic mutations have been linked to stress adaptation in *C. reinhardtii*, although these studies have not utilized MEE. For example, transposable elements (TEs) have been implicated in stress tolerance. (Kim et al., 2006) found that TEs were responsible for a significant proportion of spontaneous mutations conferring resistance to methylammonium. Similarly, in *Chlamydomonas acidophila*, an extremophile species tolerant to cadmium, researchers used transcriptomics and qRT-PCR to demonstrate that high cadmium concentrations triggered increased transposon expression and activation of genes involved in oil biosynthesis. These findings suggest a potential link between metal stress and transposon activity. Furthermore, (Nguyen et al., 2013) identified a TOC1 transposon insertion upstream of the *CrFAD7* locus, resulting in a 65% reduction in total ω-3 fatty acids and improved photosynthetic activity under heat stress. These findings underline the importance of TEs in stress adaptation and highlight the potential for further exploration of their role in herbicide resistance.

Regarding herbicides specifically, molecular studies have begun to investigate the responses of *C. reinhardtii* to paraquat, diuron and norflurazon exposure from an ecotoxicological perspective. These studies first examined growth and physiological biomarkers (Nestler et al., 2012a) and subsequently analysed proteomic changes (Nestler et al., 2012b) in response to herbicide exposure. The authors argue that the insights gained from these studies, particularly the identification of protein markers, demonstrate that proteomic profiling is a highly sensitive tool for ecotoxicological research (Nestler et al., 2012b). From an evolutionary perspective, uncovering the genetic basis of adaptation to herbicides is equally critical. Evidence of proteomic signatures linked to herbicide exposure provides a compelling rationale for integrating MEE with NGS to uncover herbicide resistance associated genomic changes.

Given the availability of the sequenced genome of *C. reinhardtii* (Merchant et al., 2007), evolve-and-resequence studies represent a powerful opportunity to uncover the genetic mechanisms underlying herbicide resistance. Combining MEE and NGS will allow to bridge the gap between phenotype and genotype, providing valuable insights into the evolution of herbicide resistance.

1.4 Aims of this thesis

This thesis explores evolutionary responses to different rates of anthropogenic environmental change, using herbicide exposure as a model context. Specifically, it investigates how varying rates of glyphosate application influence the evolution of resistance in the unicellular green alga *C. reinhardtii*, an ecologically relevant model species. While recognising that herbicide resistance presents a pressing real-world challenge, the primary motivation for this work lies in understanding the evolutionary processes that drive adaptation under changing selective conditions.

By integrating experimental evolution, fitness assays, and genomic analyses, this research examines how different selective regimes shape the dynamics, outcomes, and repeatability of resistance evolution. In addition to contributing to our understanding of adaptation under varying environmental pressures, the genomic component of this work provides novel insights into the molecular basis of glyphosate resistance in *C. reinhardtii*—a relatively underexplored area. By linking changes in phenotype with changes in genotype under selection, this thesis represents an important step forward in utilising this model organism for studying evolutionary responses to strong anthropogenic herbicide selection.

In this thesis, I attempt to address the following questions of importance to the study of evolution and adaptation:

Does exposure of *C. reinhardtii* to a simplified glyphosate selective regime (single lethal dose) lead to characteristic phenotypic adaptation, associated to measurable changes in genotype?

Does alteration of the selective regime through varying rates of glyphosate dose-increase influence the evolutionary dynamics, resistance outcomes, and associated genomic adaptations in *C. reinhardtii*?

To what extent is the evolution of glyphosate resistance in *C. reinhardtii* repeatable across replicate populations and between different glyphosate selective regimes?

What fitness costs, if any, are associated with evolved glyphosate resistance in *C. reinhardtii* under different glyphosate selection regimes?

2 Material and methods

2.1 Biological material and growth conditions

Chlamydomonas reinhardtii cultures routinely maintained in our laboratory (CC-1690 wild-type mt+ [Sager 21 gr] strain) were obtained from https://www.chlamycollection.org/. Axenic cultures of this single mating type were kept in 20 ml of Bold's media (BM), ensuring phototrophic growth and asexual cell division (Harris, 2008). *C. reinhardtii* has a short life cycle of 7-10 hours under optimal growth conditions and large populations can be contained in liquid media (Harris, 2008a).

Cultures could be stored for up to 6 months on Agar slopes (BM with additional 2 $g.l^{-1}$ Yeast extract, 1.2 g $.^{l-1}$ Na acetate and 15 $g.l^{-1}$ of agar). Each slope is inoculated with 100 µl of mature *C. reinhardtii* culture (7 to 10 days old) and placed in an incubator at 28˚C and constant LED lights (160 $\mu mol.m^{-2}. s^{-1}$) for seven days to establish the cells on the slope. After the incubation period, the slopes are stored upside down in dim light and room temperature. They can subsequently be used to inoculate liquid cultures for experiments or to transfer to fresh slants.

During experiments, cultures were maintained by weekly transfer into fresh BM and kept in a shaking incubator at 180 rpm, 28˚C and constant LED lights (160 $\mu mol.m^{-2}. s^{-1}$).

A preliminary study pre-dating this PhD (data not shown) established that there were environmental gradients in the incubator that resulted in differential growth of *C. reinhardtii* cultures. Therefore, the incubator is divided into two blocks of uneven size. For the Single Rate (Chapter 3) and the Variable Rate (Chapter 4) selection the dose-response and selection experiments were both conducted in the same single block and therefore the positions of the tubes within this homogeneous location follow a complete randomised design. During the fitness and level of resistance assays following up from the Variable Rate experiment, the whole incubator is used, hence the cultures are arranged in an incomplete randomised block design.

Prior to any experiment, contamination checks on the cultures are performed by transferring 10 µl of culture on three replicates of 2.8% w/v nutrient agar (Sigma- Aldrich) plates under sterile conditions. After 4 days at 28˚C in the dark, contamination is assessed by placing the nutrient agar plates under a binocular microscope (Leica MZ6) to confirm the absence of bacterial or fungal growth.

2.2 Estimation of Population cell density

Population growth is an estimate of fitness in the ancestral and selective environment and is inferred from estimates of population cell density ($cell.ml^{-1}$). To do this, three repeated measures of optical density by measuring absorbance at 750 nm (OD_{750}) per culture tube were performed using a spectrophotometer (Jenway 6300). Work pre-dating this PhD (data not shown) established that the detection limit of this system is of OD_{750}=0.025 (approximately 26,000 $cells.ml^{-1}$).

OD_{750} measurements of a series of dilutions (at 5, 10, 20, 30, 40, 50, 60, 80, 90 and 100% of the final volume) and cell counts were performed on three independent populations to produce a calibration equation allowing to estimate the cell concentration of *C. reinhardtii* populations. Three 500µl aliquots of the 100 dilutions were set aside in the dark at 4°C to stop cell division. To perform cell

counts, 10µl of an additional quarter dilution in ddH_2O of each aliquot was placed on a 0.1ul haemocytometer grid. Counts for each aliquot was replicated three times. The cells were counted under a stereomicroscope (M205 FA). The cell concentration per ml was calculated by multiplying the average number of cells by the dilution factor and 10^4. The OD_{750} and cell count pairs were used to determine the relationship between optical density and cell concentration by finding the curve of best fit (data not shown for the first calibration as it pre-dates this PhD).

A different calibration was performed for the SR and VR selection experiment. The variance accounted for the linear models used for calibration was of 94% in the case of the SR selection experiment, and of 97.4% in the case of the VR selection experiment. The resulting quadratic function to estimate cell concentration are:

- Single Rate experiment: Number of cells $.ml^{-1}$ = 548,069 OD_{750} + 559,257 OD_{750}^2
- Variable Rate experiment: Number of cells $.ml^{-1}$ = 1,046,213 OD_{750} + 875,730 OD_{750}^2

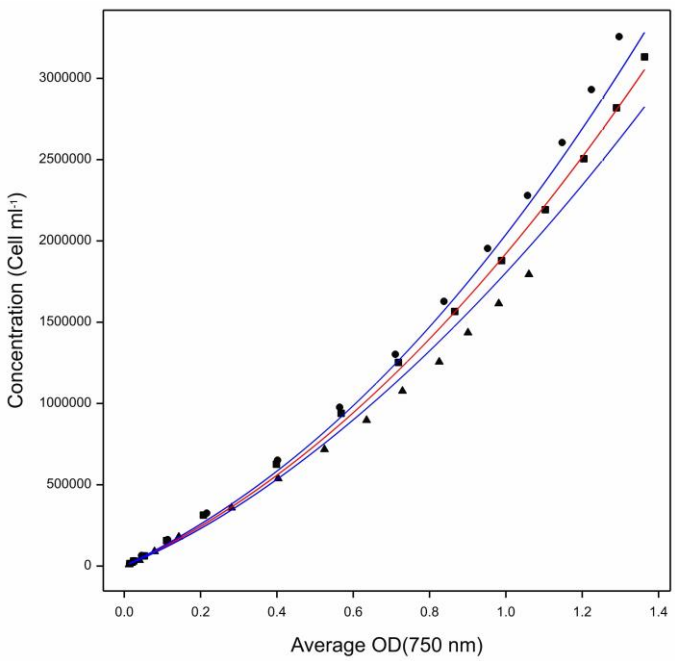

Figure 4: Cell concentration-optical density relationship from the Variable Rate experiment. Measurements were taken from three independent 7-day-old *C. reinhardtii* cultures (distinguished by symbols). OD_{750} was measured at a range of dilutions (5,10,20,30,40,50,60,80,90 and 100% of the final volume). Cells counted in triplicates on three aliquots at 100% dilutions and converted to concentrations for each culture. The cell concentrations at the other dilutions were estimated by multiplying the cell concentration at 100% by the corresponding dilution factor. The relationship was modelled using a weighted quadratic (polynomial of order 2) regression with the intercept constrained to be zero and weights equal to the reciprocal of the squared cell concentrations to achieve variance homogeneity. The red line is the best fit curve, and the blue lines are 95% confidence intervals around the fitted line.

2.3 Source-Sink Scenario for selection experiments:

Both selection experiments (Chapter 3 and Chapter 4) used a source-sink scenario (Kawecki and Holt, 2002) as previous work demonstrated (Lagator et al., 2012, 2013, 2014) that this allowed selected populations to adapt to strong and sudden environmental change (glyphosate at 1 MIC)

through sustained immigration from their source population. The source-sink scenario corresponds to immigration from an unselected population (large population) into a glyphosate selected population (small population). It can be seen as a metapopulation with unidirectional immigration. On the first day of selection, tubes were inoculated with 125,000 cells form an isogenic population. Replicate source-sink lines were established in each of the environments as described in Figure 5. Population cell density was estimated after seven days (exponential growth phase) prior to the weekly 1% volume serial cell transfer into fresh media. A minimum cell density of 625,000 cells.ml^{-1} (which corresponds to a minimum of 6.64 cell divisions in seven days) was the threshold over which a population would avoid being driven to extinction due to weekly bottlenecks at serial transfer. If cell density was lower than 625,000 cells .ml^{-1}, as expected until resistance would develop in a selected population, the appropriate replicate source population would provide additional cells to make the total inoculum at 125,000 cells (Lagator et al., 2012, 2013, 2014).

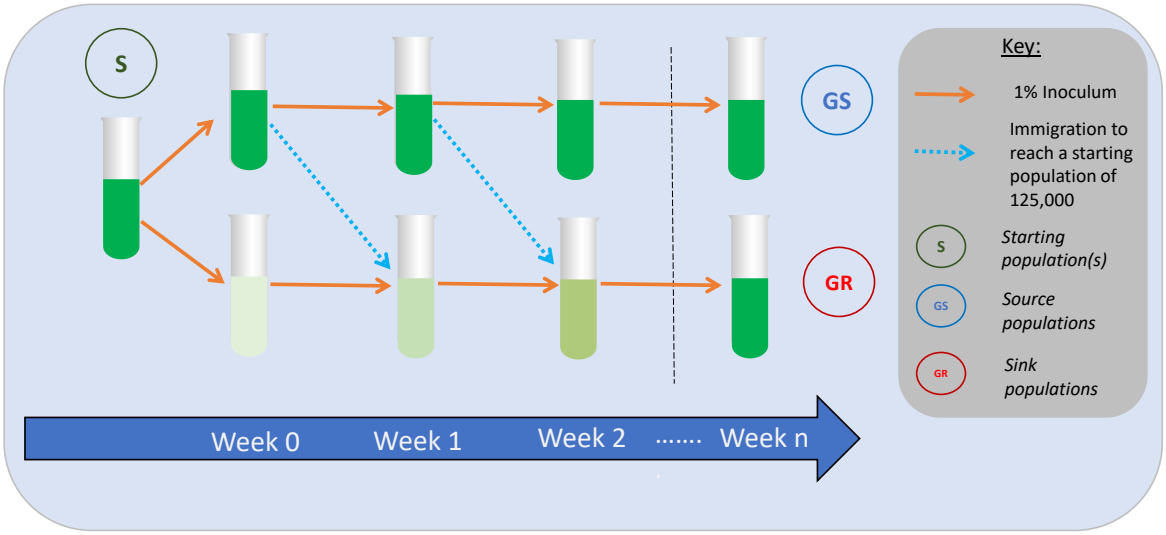

Figure 5: Source-sink population structure used during Glyphosate selection experiments. Cultures reproduced asexually for the whole duration of the experiments described in this thesis. Independent replicate source-sink populations were established from glyphosate sensitive (GS) population. While five distinct lines were used to start the Single Rate experiment, the design was later improved by using a single isogenic starting population in the Variable Rate experiment. Populations are maintained by a weekly 1% serial transfer in fresh media. Sink populations with low cell densities receive additional cells from their respective source population to ensure a minimum starting population to avoid extinction due to bottleneck effects. Immigration is stopped when these populations' growth is high enough to maintain themselves after serial transfer (i.e when 200 µl of culture contains at least 125,000 cells).

2.3.1 Glyphosate Stocks

A master glyphosate stock solution (sourced from Sigma Aldrich, purity 99%) for each selection experiment were made directly in BM, filter sterilised and kept at 4°C for the duration of the experiment. Desired concentrations of glyphosate selective media was prepared for weekly transfer by adding 2 ml of 10X concentrated solutions into 18 ml of sterile BM.

2.4 Preparation for Whole genome sequencing

2.4.1 Isogenic populations for DNA extraction

Isogenic population derived from the SR and VR selection experiment were produced by streaking cultures on BM agar plates under sterile conditions. The agar plates are placed in an incubator at 28˚C and constant LED lights (160 µmol.m^{-2}. s^{-1}) for 3 to 4 days. Single colony growth is checked under a binocular microscope and selected single colonies are delimited on the back of the petri dish with a marker pen. Under sterile conditions, the single colonies are transferred with a loop in 100 ml of BM (GS) or BM containing glyphosate at 1MIC (GR) and placed in a shaking incubator at 180 rpm, 28˚C and constant LED lights (160 µmol.m^{-2}. s^{-1}) to divide until the isogenic culture contained a minimum of 60 million cell. Before cell harvest, the isogenic culture is checked for contamination by transferring 10 µl of culture on three replicates of 2.8% w/v nutrient agar (Sigma- Aldrich) plates under sterile conditions. After 4 days at 28˚C in the dark, contamination is assessed.

2.4.2 Cell harvest

Cell harvest for DNA extraction was performed on 100ml isogenic axenic cultures of a minimum 60 million cells to yield approximately 3µg of DNA after extraction. The isogenic cultures were divided in two 50ml centrifuge tubes and centrifuged at 5500g for 10 minutes to form a cell pellet. Cells were resuspended in 1ml of sterile dH_2O and transferred into 2ml sterile microcentrifuge tubes for centrifuging at max speed for 5 mins to form the final cell pellet for DNA extractions. The supernatant was discarded without disturbing the pellet and the tubes were kept at -80°C for long term storage.

2.4.3 DNA extractions

DNA extractions were performed using a modified in-house leaf DNA extraction method adapted for Chlamydomonas cells (for the detailed protocol, see Appendix 2), the main modification being in the cell lysis step. This involved performing three snap freezing in liquid nitrogen (2 mins) and thawing (2 mins) cycles before adding 1ml of fresh genomic extraction buffer and a glass bead to homogenize in with a bead beater at max speed for 20 seconds. The rest of the leaf extraction method was left unchanged. Nucleic acids were eluted in TER and incubated at 50°C for 1 hour with gentle mixing every 15 mins to digest RNA. A final centrifuge step of 5 mins at 16,000 rcf was added to remove polysaccharides. DNA extraction quality control was performed by agarose gel electrophoresis (1% agarose gel run at 100V for 90 mins with, 100ng of samples run alongside an Invitrogen 1KB+ Ladder) to assess genomic DNA integrity. A Nanodrop ND-1000 spectrophotometer (Thermo Scientific) was used to assess the presence of extraction contaminants. Sample concentration was determined using a Qubit dsDNA BR Assay Kit (Life Technologies) as per the manufacturer's recommendation. All

samples tested passed the QC requirements for sequencing at the Earlham Institute (SR samples) and The Center for Genomic Research, University of Liverpool (VR samples).

2.5 Bioinformatics analysis

2.5.1 Variant calling pipelines

2.5.1.1 Reference genome

C. reinhardtii genome is GC rich (64%) and non-compact (110Mb) with genes carrying on average 7 introns (of >350 bp) (Merchant et al., 2007), and relatively active transposable elements (Kim et al., 2006).

There are now several reference genomes published for *C. reinhardtii.* The first *C. reinhardtii* genome was assembled from sequencing a CC-503 strain and published by Merchant et al (2007), with the latest version (V5.6) available from the JGI Phytozome13 portal. It is subsequently referred to as Phytozome-CC503_V5.6 in this thesis. The Phytozome-CC503_V5.6 *C. reinhardtii* genome is contained in 54 scaffolds (17 chromosomes and 37 minor scaffolds) assembled from 1,495 contigs (Merchant et al., 2007)If~2 Mb of sequence remains unplaced,this genome remains the most advanced in terms of annotation.

Recent advances in sequencing technologies offer the prospect of generating high quality reference assemblies for a greater number of strains with relative ease . This may help circumvent problems that may arise when using a particular genome reference for analysis of experimental data generated with a differnet strain (Payne et al., 2023). Though CC-1690 (used in this study) and CC-503 strains are relatively closely related and most of the genome is identical by descent, a previous study by Flowers et al. (2015) reported significant levels of polymorphism. The genome of CC1690 features large duplications on chromosome 13, 12 new candidate genes and ~61,480 SNP when compared to the Phytozome-CC503_V5.6.

More recently, a highly contiguous nanopore nuclear genome assembly of our laboratory strain CC-1690 was published (O'Donnell et al., 2020).The assembly has a total length of 111Mb, is composed of 21 contigs ,18 chromosomes and contains five more complete benchmarking universal single-copy orthologs (BUSCOs) than the Phytozome-CC503_V5.6 reference (O'Donnell et al., 2020). However, no annotation is publicly available for this genome. A comparison of variant calling results obtained with when both references were used is provided in Chapter 3. The O'Donnell-CC1690_v1.0 reference genome produced the best results and was thus used for all subsequent alignment, variant calling, and genotyping analysis on samples from both Single Rate (Chapter 3) and Variable Rate (Chapter 5) experiments. For the purpose of the studies described in this thesis, the assembly was modified to include chloroplast and mitochondrial genome sequences from CC-503cm92mt+ genome version 3.1 (archived on Phytozome https://phytozome-next.jgi.doe.gov/) to help avoid misalignment of reads derived from these organelle genomes to the nuclear genome (Ness et al., 2015). In the absence of a dedicated annotation, the Phytozome-CC503_V5.6 annotation was transferred to the O'Donnell-CC1690_v1.0 by leveraging sequence similarity and structural synteny, mapping individual genes and preserving exon boundaries and handling structural rearangements or gene duplications with with Liftoff(v1.6.3) (Shumate and Salzberg, 2021).

2.5.1.2 Sequence alignement and processing

Pre-processing was performed on the Single Rate (SR) raw reads (Chapter 3) with Trimmomatic v0.39 (Bolger et al., 2014) while Variable rate (VR) reads (Chapter 5) were pre-processed by the Center for Genomic Research prior to sequence delivery. Raw fastq files were trimmed to exclude Illumina TruSeq2-PE adapter sequences and reads ends were then further trimmed with a minimum window size of 5 and quality score of 20. Reads shorter than 50 bp after trimming were removed. FastQC was used for read quality control. This tool produces an overview of key parameters such as read quality, read length and GC content (Andrews, 2010). Read QC was performed both on raw and trimmed reads with FastQC v0.11.9.

High quality paired-end reads were subsequently aligned with the Burrows-Wheeler Aligner (BWA-MEM v0.7.17) with default settings. BWA-MEM's algorithm chooses automatically between local and end-to-end alignments (Li, 2013). Alignment quality control was assessed with Qualimap v2.2.1 using the *bamqc* command to generate comprehensive quality reports (García-Alcalde et al., 2012). Aligned bam files were further processed by sorting them by coordinates with SAMtools v1.18 (Li et al., 2009) and read groups were added with Picard tools v3.0.0 (Broad Institute, 2019) in preparation for variant calling.

2.5.1.3 Variant calling and genotyping

Variant calling was performed using Bayesian genetic variant detection with Freebayes v 1.3.6 configured for haploid genomes. Unlike alignment-based variant callers, Freebayes limits problems caused by sequences having multiple possible alignments (Garrison and Marth, 2012). Considering the *C. reinhardtii* genome is known to contain many repeats and transposons (Vallon and Dutcher, 2008) Freebayes was deemed to be suitable. Alignments with mapping quality below 20 and reads with supporting base quality below 15 were excluded from the analysis. For each of sample, a variant was only evaluated if a minimum of ten alternative alleles were found. To avoid calling false variants, Freebayes performs local realignment around indels. To ensure that the same local realignment solutions were chosen, variants were called simultaneously on all samples within an experiment (Ness et al., 2012).

2.5.1.4 Filtering of variant list to identify loci potentially linked to evolved glyphosate resistance

To keep calls of interest to our study (i.e. calls in loci potentially linked to an observed glyphosate-resistant phenotype), VCF files where further processed in R (v4.2.3) (R Core Team, 2023) with the package vcfR v1.14.0 (Knaus and Grünwald, 2017) and further filtered to produce a list of candidate genes potentially involved in *C. reinhardtii* glyphosate resistance in the context of the SR and VR selection experiments (Chapter 3 and Chapter 4 & 5 respectively). Filters were applied to retain variants according to the following criteria:

- Variants for which, within a replicate GS/GR pair, the genotype differed between the GR and GS sample.
- Variants for which the GR allele is not the reference allele.
- Variants for which the GR allele is not found in any of the GS samples.

To retain only high-quality variants supporting glyphosate resistance, we kept loci for which alternate allele call was supported by high fraction of GR sample's reads. Because samples are haploid, the minimum alternate allele frequency was set to 0.85 for high confidence variant calling. The alternate allele frequency (AAF) was calculated as follow:

$$AAF = \frac{AO}{AO + RO}$$

- AO: number of reads supporting alternate allele
- RO: number of reads supporting reference allele
- AO+RO: the total number of reads for a biallelic loci

Visual confirmation of variants to refine the variant calling pipeline was performed with Geneious v10.2.3 (Kearse et al., 2012) by comparing reads from both GR and GS samples aligned to the reference genome at the variant locus.

Box 1: The Variant calling analysis pipeline used for Single Rate (Chapter 3) and Variable Rate (Chapter 5) experiments: steps and tools used.GR refers to glyphosate resistant samples and GS refers to glyphosate sensitive samples.

A. Paired-end reads QC and processing: Keep high quality reads for downstream analysis

- Raw paired-end read quality check (FastQC v0.11.9): assess the quality of raw paired-end reads after sequencing.
- Read trimming (Trimmomatic v0.39): keep high-quality, adapter free reads:
  - Adapter trimming: removes any residual Illumina TruSeq2-PE adapter sequences allowing 2 mismatches on a 10 bp window size and a minimum score of 30.
  - Quality trimming: uses a 5 bp sliding window to trim low-quality ends of reads with a minimum quality score of 20.
  - Minimum read length filtering removes very short reads (50 bp) and reduces misalignment.
- Trimmed paired-end read quality check (FastQC v0.11.9): assess the quality of trimmed paired-end reads to verify read quality has been improved and data is suitable for downstream analysis.

B. Alignment and QC: map high quality reads on a reference genome and assess mapping quality prior to downstream analysis

- Published reference genome from our laboratory strain: O'Donnell-CC1690_v1.0.
- Map paired-end trimmed reads to a reference genome (BWA-MEM v0.7.17): with aligner's default parameters.
- Process SAM files into sorted and indexed BAM files for variant calling analysis (SAMtools v1.18):
  - Converts SAM files into compact BAM files for faster processing.
  - Sorts the BAM file by alignment position in the reference genome.
  - Index the sorted BAM file enabling rapid random access to specific regions of the file.
- Add essential metadata to BAM files (Picard Tools v3.0.0 AddOrReplaceReadGroups): Adds read group information to identify and differentiate samples and create indexed BAM files for downstream analysis.
- Perform quality control analysis on BAM files (Qualimap v2.2.1 bamqc): Generates a comprehensive report on alignment data and providing insights on both quality and characteristics of the sequencing data and read alignments to the reference genome.

C. Variant calling analysis tailored to our datasets

- Variant calling: Identify potential genetic variants -SNPs and Indels (Freebayes v1.3.6): bayesian genetic variant detection from aligned reads with the following settings:
  - Call variants on all samples simultaneously to ensure same local realignments around indels are performed.
  - Reads with a base quality score below 15 are ignored to reduce noise by increasing base call accuracy.
  - Read mapping quality of 20 or greater to be included in analysis and calling variants in reads for which there is high confidence in the alignment to reference.
  - Setting ploidy of 1 for the *C. reinhardtii* haploid genome.
  - Minimum of 10 reads supporting the alternate allele required to make a variant call to discard potentially erroneous calls in low depth regions.
- Variant Annotation and effect prediction (SnpEff v4.3+T.galaxy2): informs on variant location (gene name and coding region, intron, intergenic) and predicts the impact of the variant on the protein (synonymous changes, missense mutations, nonsense mutations, or frameshifts).

D. Post-variant calling filtering (R v4.2.3): keep calls that support glyphosate resistance in each GR/GS replicate

- Keep calls at positions where genotypes differ between GR and GS at the replicate level (i.e. $GR_a \neq GS_a$).
- Keep calls where GR genotype is the not the reference (which is GS) allele (i.e. GRa ≠ 0).
- Keep calls when GR genotype is not present in any of the GS samples (i.e. $GR_a \neq GS_b \neq GS_c$).
- Keep calls when GR alternate allele frequency is at least 0.85 (i.e.: AAF for GR_i is ≥ 0.85).

2.5.2 Variant annotation and description

Variant annotation and effect prediction was performed using SnpEff v4.3+Tgalaxy2 on the Galaxy EU platform with the *C. reinhardtii* O'Donnell-CC1690_v1.0 annotation (unpublished). SnpEff is a tool that categorises (Table 2), annotates and predicts the effects (Table 3) of genetic variants (Cingolani et al., 2012a). SnpEff also assigns impact scores (High, moderate, low and modifier) to genetic variants by evaluating their potential effects on gene function based on variant type, location and predicted effect on the gene product (Table 4). Like any prediction algorithm, SnpEff predictions must be used with care and require validation by wet-lab experiments.

Table 2:Detailed Variant type categories in SnpEff (adapted from https://pcingola.github.io/SnpEff/snpeff/introduction/ accessed 09/01/2025)

Type		Example	
		Reference:	**Sample:**
SNP	(Single-Nucleotide Polymorphism)	A	C
Ins	(Insertion)	A	AGT
Del	(Deletion)	AC	C
MNP	(Multiple-nucleotide polymorphism)	ATA	GTC
MIXED	(Multiple-nucleotide and an InDel)	ATA	GTCAGT

Table 3: Detailed effect list from SnpEff adapted from Cingolani et al.(2012).

Effect	Note
CDS	The variant hits a CDS
CODON_CHANGE	One or many codons are changed
CODON_CHANGE_PLUS_CODON_DELETION	One codon is changed and one or more codons are deleted
CODON_CHANGE_PLUS_CODON_INSERTION	One codon is changed and one or many codons are inserted
CODON_DELETION	One or many codons are deleted
CODON_INSERTION	One or many codons are inserted
DOWNSTREAM	Downstream of a gene (default length: 5K bases)
EXON	The variant hits an exon
EXON_DELETED	A deletion removes the whole exon.
FRAME_SHIFT	Insertion or deletion causes a frame shift
GENE	The variant hits a gene
INTERGENIC	The variant is in an intergenic region
INTERGENIC_CONSERVED	The variant is in a highly conserved intergenic region

INTRON	Variant hist and intron. Technically, hits no exon in the transcript
INTRON_CONSERVED	The variant is in a highly conserved intronic region
NON_SYNONYMOUS_CODING	Variant causes a codon that produces a different amino acid
SPLICE_SITE_ACCEPTOR	The variant hits a splice acceptor site (defined as two bases before exon start, except for the first exon)
SPLICE_SITE_DONOR	The variant hits a Splice donor site (defined as two bases after coding exon end, except for the last exon)
START_GAINED	A variant in 5'UTR region produces a three base sequence that can be a START codon
START_LOST	Variant causes start codon to be mutated into a non-start codon
STOP_GAINED	Variant causes a STOP codon
STOP_LOST	Variant causes stop codon to be mutated into a non-stop codon
SYNONYMOUS_CODING	Variant causes a codon that produces the same amino acid
SYNONYMOUS_START	Variant causes start codon to be mutated into another start codon
SYNONYMOUS_STOP	Variant causes stop codon to be mutated into another stop codon
TRANSCRIPT	The variant hits a transcript
UPSTREAM	Upstream of a gene (default length: 5K bases)
UTR_3_DELETED	The variant deletes an exon which is in the 3'UTR of the transcript
UTR_3_PRIME	Variant hits 3'UTR region
UTR_5_DELETED	The variant deletes an exon which is in the 5'UTR of the transcript
UTR_5_PRIME	Variant hits 5'UTR region

Table 4: SnpEff details on impact score annotations (adapted from https://pcingola.github.io/SnpEff/snpeff/inputoutput/ accessed on the 09/01/2025). High impact variants are likely to cause significant disruptions to the gene product, Moderate impact variants may change protein function, Low impact variants are expected to have minimal effects and Modifier variants are mainly located in non-coding regions and have uncertain predicted impacts.

Impact scores	Snpeff effect	Description
HIGH	CHROMOSOME_LARGE_DELETION	A large part (over 1%) of the chromosome was deleted
HIGH	CHROMOSOME_LARGE_DUPLICATION	Duplication of a large chromosome segment (over 1% or 1,000,000 bases)
HIGH	CHROMOSOME_LARGE_INVERSION	Inversion of a large chromosome segment (over 1% or 1,000,000 bases)
HIGH	EXON_DELETED	A deletion removes the whole exon
HIGH	EXON_DELETED_PARTIAL	Deletion affecting part of an exon
HIGH	EXON_DUPLICATION	Duplication of an exon
HIGH	EXON_DUPLICATION_PARTIAL	Duplication affecting part of an exon
HIGH	EXON_INVERSION	Inversion of an exon
HIGH	EXON_INVERSION_PARTIAL	Inversion affecting part of an exon
HIGH	FRAME_SHIFT	Insertion or deletion causes a frame shift. e.g.: An indel size is not multiple of 3
HIGH	GENE_DELETED	Deletion of a gene
HIGH	GENE_FUSION	Fusion of two genes
HIGH	GENE_FUSION_HALF	Fusion of one gene and an intergenic region
HIGH	GENE_FUSION_REVERSE	Fusion of two genes in opposite directions
HIGH	GENE_REARRANGEMENT	Rearrangement affecting one or

		more genes
HIGH	PROTEIN_PROTEIN_INTERACTION_LOCUS	Protein-Protein interaction loci
HIGH	PROTEIN_STRUCTURAL_INTERACTION_LOCUS	Within protein interaction loci (e.g. two AA that are in contact within the same protein, possibly helping structural conformation)
HIGH	RARE_AMINO_ACID	The variant hits a rare amino acid thus is likely to produce protein loss of function
HIGH	SPLICE_SITE_ACCEPTOR	The variant hits a splice acceptor site (defined as two bases before exon start, except for the first exon)
HIGH	SPLICE_SITE_DONOR	The variant hits a Splice donor site (defined as two bases after coding exon end, except for the last exon).
HIGH	START_LOST	Variant causes start codon to be mutated into a non-start codon. e.g.: aTg/aGg, M/R
HIGH	STOP_GAINED	Variant causes a STOP codon. e.g.: Cag/Tag, Q/*
HIGH	STOP_LOST	Variant causes stop codon to be mutated into a non-stop codon. e.g.: Tga/Cga, */R
HIGH	TRANSCRIPT_DELETED	Deletion of a transcript
MODERATE	CODON_CHANGE_PLUS CODON_DELETION	One codon is changed and one or more codons are deleted (e.g.: A deletion of size multiple of three, not at codon boundary)

MODERATE	CODON_CHANGE_PLUS CODON_INSERTION	One codon is changed and one or many codons are inserted. e.g.: An insert of size multiple of three, not at codon boundary
MODERATE	CODON_DELETION	One or many codons are deleted (e.g.: A deletion multiple of three at codon boundary)
MODERATE	CODON_INSERTION	One or many codons are inserted (e.g.: An insert multiple of three in a codon boundary)
MODERATE	NEXT_PROT	A 'NextProt' based annotation. Details are provided in the 'feature type' sub-field (ANN), or in the effect details (EFF)
MODERATE	NON_SYNONYMOUS_CODING	Variant causes a codon that produces a different amino acid (e.g.: Tgg/Cgg, W/R)
MODERATE	SPLICE_SITE_BRANCH_U12	A variant affective putative (Lariat) branch point from U12 splicing machinery, located in the intron
MODERATE	UTR_3_DELETED	The variant deletes an exon which is in the 3'UTR of the transcript
MODERATE	UTR_5_DELETED	The variant deletes an exon which is in the 5'UTR of the transcript
LOW	CODON_CHANGE	One or many codons are changed (e.g.: An MNP of size multiple of 3)

LOW	NON_SYNONYMOUS_START	Variant causes start codon to be mutated into another start codon (the new codon produces a different AA) e.g.: Atg/Ctg, M/L (ATG and CTG can be START codons)
LOW	NON_SYNONYMOUS_STOP	Variant causes stop codon to be mutated into another stop codon (the new codon produces a different AA). e.g.: Atg/Ctg, M/L (ATG and CTG can be START codons)
LOW	SPLICE_SITE_BRANCH	A variant affective putative (Lariat) branch point, located in the intron
LOW	SPLICE_SITE_REGION	A sequence variant in which a change has occurred within the region of the splice site, either within 1-3 bases of the exon or 3-8 bases of the intron
LOW	START_GAINED	A variant in 5'UTR region produces a three base sequence that can be a START codon
LOW	SYNONYMOUS_CODING	Variant causes a codon that produces the same amino acid. e.g.: Ttg/Ctg, L/L
LOW	SYNONYMOUS_START	Variant causes start codon to be mutated into another start codon. e.g.: Ttg/Ctg, L/L (TTG and CTG can be START codons)

LOW	SYNONYMOUS_STOP	Variant causes stop codon to be mutated into another stop codon. e.g.: taA/taG, */*
MODIFIER	CDS	The variant hits a CDS
MODIFIER	DOWNSTREAM	Downstream of a gene (default length: 5K bases)
MODIFIER	EXON	The variant hits an exon (from a non-coding transcript) or a retained intron
MODIFIER	GENE	The variant hits a gene.
MODIFIER	GENE_DUPLICATION	Duplication of a gene.
MODIFIER	INTERGENIC	The variant is in an intergenic region
MODIFIER	INTERGENIC_CONSERVED	The variant is in a highly conserved intergenic region
MODIFIER	INTRAGENIC	The variant hits a gene, but no transcripts within the gene
MODIFIER	INTRON	Variant hits and intron. Technically, hits no exon in the transcript
MODIFIER	INTRON_CONSERVED	The variant is in a highly conserved intronic region
MODIFIER	MICRO_RNA	Variant affects a miRNA
MODIFIER	REGULATION	The variant hits a known regulatory feature (non-coding)
MODIFIER	TRANSCRIPT	The variant hits a transcript
MODIFIER	UPSTREAM	Upstream of a gene (default length: 5K bases)
MODIFIER	UTR_3_PRIME	Variant hits 3'UTR region
MODIFIER	UTR_5_PRIME	Variant hits 5'UTR region

2.5.3 Gene Ontology analysis

High throughput experimental techniques can produce huge quantities of data and therefore there is a need to develop techniques to capture biological information. Investigating shared functions among genes can be achieved using the biological knowledge provided by biological ontologies such as Gene Ontology (GO). In GO, gene functions are described in three distinct aspects: a gene encodes a gene product which carries out a molecular level activity (named Molecular Function) in a specific location (named Cellular Component) and this activity contributes to a larger biological objective (named Biological Process) (Thomas, 2017). GO terms are standardised capturing biological knowledge in these three formalized ontologies (Biological Process, Molecular Function, and Cellular Component).

Typically, extensive lists of differentially expressed genes are examined using Gene Ontology (GO) term enrichment analysis, a computational approach designed to determine whether specific GO terms are statistically over-represented within a given gene list, thereby providing insights into the underlying biological processes. In this thesis, however, focus was given to identifying patterns within a gene list of interest generated by variant calling analysis. This dataset did not yield a sufficient number of candidate genes to achieve the statistical power required for GO term enrichment analysis, even after combining variant calling data from both SR and VR experiments. To summarise these datasets, heatmaps of GO terms were generated for each of the three ontologies and presented in Chapter 5.

3 Investigating the genomic basis of glyphosate adaptation in *C. reinhardtii*

3.1 Introduction

Modern agriculture has relied mainly on chemicals for weed control which enabled the evolution of resistance to herbicides. It is now a major threat to food security in the current agricultural model. Glyphosate was introduced as a new compound in 1974. Its use became widespread as evolution of resistance to this broad-spectrum herbicide was considered to be unlikely (Powles, 2008). Inevitably, the first case of a glyphosate-resistant weed was found in 1998. To date more than 60 weed species have evolved glyphosate resistance (Heap, 2021).

Glyphosate inhibits reactions of the shikimate pathway leading to synthesis of aromatic amino acids in plants, fungi, bacteria and yeast (Bentley, 1990). Glyphosate resistance has been studied extensively in plants (Baek et al., 2021; Bradshaw et al., 1997; Délye and Christophe, 2013; Sammons and Gaines, 2014) and to a lesser extent in microbes (Hertel et al., 2021; Hove-Jensen et al., 2014; Patriarcheas et al., 2023; Pollegioni et al., 2011). Glyphosate targets the EPSPS (in plants) and AroA (in microbes) enzyme that catalyses transfer of enolpyruvyl moiety of phosphoenolpyruvate (PEP) to the 5-hydroxyl of shikimate 3-phosphate (S3P) to produce 5-enolpyruvyl shikimate 6-phosphate (EPSP) and inorganic phosphate (Patriarcheas et al., 2023). Glyphosate resistance mechanisms are generally categorised as target-site resisance (TSR) when they involve modifications to glyphosate target enzyme ecoded *EPSPS* or *AroA/Aro1* genes, and non-target-site resistance (NTSR) when encompassing mechanisms excluding mutations in these target genes. Several of these mechanisms are found in plants, bacteria, yeast and fungi. Glyphosate TSR mechanisms uncovered to date involve mutations, gene copy number variation (GCNV), overexpression and increased enzyme activity, all of which have been described in weeds. In microbes, only mutations (Liu and Cao, 2018; Pollegioni et al., 2011) and GCNV (Ravishankar et al., 2020b) have been reported. NTSR mechanisms are complex multigenic traits involved in reduced glyphosate uptake, increased translocation (in plants)/efflux (in microbes), enhanced metabolic degradation and detoxification. In yeast, a recent study found polymorphisms in the coding region of 148 NTSR genes and evidence for GCNV potentially due to Ty transposable elements (Ravishankar et al., 2020b).

Herbicide resistance is the result of human-driven evolution and requires to take an evolutionary approach to gain fundamental understanding of its mechanisms (Neve et al., 2009). Indeed, taking an experimental evolutionary approach allows the observation of evolution in action: different selection pressure in time and space can be applied under controlled conditions on large sensitive populations over several generations (Buckling et al., 2009; Neve et al., 2009; Sammons and Gaines, 2014). One of the most promising approaches is experimental evolution with model organisms, due to their large population size and short generation time, which favour rapid evolution. Additionally, most of them have simple genomes that have already been sequenced allowing investigation of the genetic basis of adaptation (Bell and Reboud, 1997; Buckling et al., 2009; Elena and Lenski, 2003).

Chlamydomonas reinhardtii is a perfect candidate to conduct experimental evolution of herbicide resistance studies. It has been used as a model species in experimental evolution since the 90s (Bell, 1990a, 1990b). This unicellular green chlorophyte reproduces mainly asexually but sexual reproduction can be induced. It has a short life cycle of 7-10 hours under optimal growth conditions and large populations can be cultured in liquid media (Harris, 2008a). It is susceptible to herbicides (Reboud, 2002) and is already a model species to study herbicide resistance (Reboud et al., 2007).

Evolution of herbicide resistance in Chlamydomonas has been studied taking an evolutionary biology approach to investigate the effect of management practices (herbicide sequential application, cycling, mixture and dose) and their underlying ecological and evolutionary theories (Lagator, 2012b; Lagator et al., 2012, 2013, 2014). In the experimental conditions tested, the evolutionary dynamics of resistance depended on the herbicides. Therefore, no universal resistance management strategies could be recommended. The major findings were that the management strategies tested often selected for generalist phenotypes and that their effect on fitness cost was unpredictable. Other studies focused on the evolution of glyphosate resistance at lethal and sub-lethal doses (Hansson et al., 2024) or increasing glyphosate doses (Melero-Jiménez et al., 2021) and associated fitness costs on which they found contrasting results. This work focused on observations of the evolved phenotype. Consequently, insight would be gained from understanding the genetic mechanisms underpinning the evolution of herbicide resistance (Lagator, 2012c).

C. reinhardtii is a powerful model organisms for genetic studies since its haploid genome has been sequenced (Merchant et al., 2007), allowing investigation of the genetic mechanisms underlying herbicide resistance. *C. reinhardtii* laboratory strains were traditionally categorised in 3 main lineages (Sager, Cambrige and Ebersold-Levine lineages), all descending from a strain isolated from a field by Smith in 1945 (Harris, 2008b). However, recent genomic comparisons of 39 strains suggest that some strains have been misidentified, and a five-lineage model (lineages I to V) provides a more accurate classification framework (Gallaher et al., 2015). Although laboratory strains have a recent common ancestry, two studies revealed notable genome differences (Flowers et al., 2015; Gallaher et al., 2015). The first versions of *C. reinhardtii* genome were assembled from sequencing the CC-503 strain from the Ebersold-Levine lineage (Merchant et al., 2007) while the CC-1690 strain recently used to uncover the mechanisms of evolution of herbicide resistance is from the Sager lineage (Hansson et al., 2024; Lagator et al., 2012, 2013, 2014b, 2014a; Vogwill et al., 2012). More recently, a genome for the CC-1690 strain has been assembled using nanopore sequencing technology (O'Donnell et al., 2020).

In this study, I chose to focus on uncovering the genetics basis of glyphosate resistance using *C. reinhardtii* CC-1690 strain. I hypothesize that whole genome re-sequencing of both sensitive (GS) and glyphosate-resistant (GR) lines combined with variant calling should inform us on the mechanisms underlying glyphosate resistance in *C. reinhardtii*. To test this, I have 3 main objectives.

First, I will (i) select for glyphosate resistance in *C. reinhardtii*, generating GR lines from GS lines for whole genome re-sequencing. I will investigate if glyphosate resistance in *C. reinhardtii* is supported by (ii) target-site resistance (TSR) and or (iii) non-target-site resistance (NTSR) mechanisms. I report that glyphosate-resistant phenotypes in Chlamydomonas arose under selection at minimum inhibitory concentration in all replicates and that results suggest the presence of NTSR mechanisms in all replicates. Although all replicates exhibit a consistent GR

phenotype, results suggest that underlying genetic basis of glyphosate resistance differ between replicates.

3.2 Methods

3.2.1 Biological material and culture conditions

C.reinhardtii cells (CC-1690 wild-type mt+ [Sager 21 gr] strain) routinely maintained in the laboratory were obtained from the Chlamydomonas Resource Center (https://www.chlamycollection.org/). Axenic cultures were kept in Bold's media (Harris, 2008a) in an Multitron Pro shaking incubator (Infors) at 28˚C, 180 rpm and constant LED light (160 $\mu mol.m^{-2}.s^{-1}$). Prior to the experiment, cultures were transferred to fresh BM each week.

A 10X solution of glyphosate (Sigma Aldrich, purity 99%) was made directly in BM and stored at 4°C for prior to use throughout the experiment for weekly preparation of selective media.

3.2.2 Single Rate (SR) selection experiment

In the SR experiment, *C. reinhardtii* populations are either kept in ancestral media or suddenly exposed to a unique glyphosate dose. Five distinct glyphosate-sensitive (GS) lines were previously maintained in Bold's medium (BM) to acclimate to culture conditions prior to the initiation of selection. Although these five lines originated from the same strain or genetic background, they were not derived from a single colony; therefore, the clonality of the five genomes cannot be presumed. On week 0, GS lines were duplicated to inoculate fresh BM to serve as controls and to inoculate BM containing glyphosate at $100mg.l^{-1}$ (Figure 6A). This dose had been established as the Minimal Inhibitory Concentration (MIC) by a dose-response experiment (Figure 6B).

Growth of the ten glyphosate selection experiment cultures was monitored by taking a weekly optical density measurements at 750 nm (OD_{750}) as an estimate for total cell biomass using a Jenway 6300 spectrophotometer. The selection experiment was set up using a source-sink experimental design on cultures undergoing weekly transfer in fresh media (Figure 6 A and see Chapter 2 for details) In the first weeks of selection, lines exposed to glyphosate needed a supplementary inoculum from the source population they originated from until growth was sufficient to ensure a starting population of 125,000 cells and avoid extinction. Glyphosate resistance was assessed when a line under selection reached a threshold biomass (625,000 cells/ ml) after 7 days of growth.

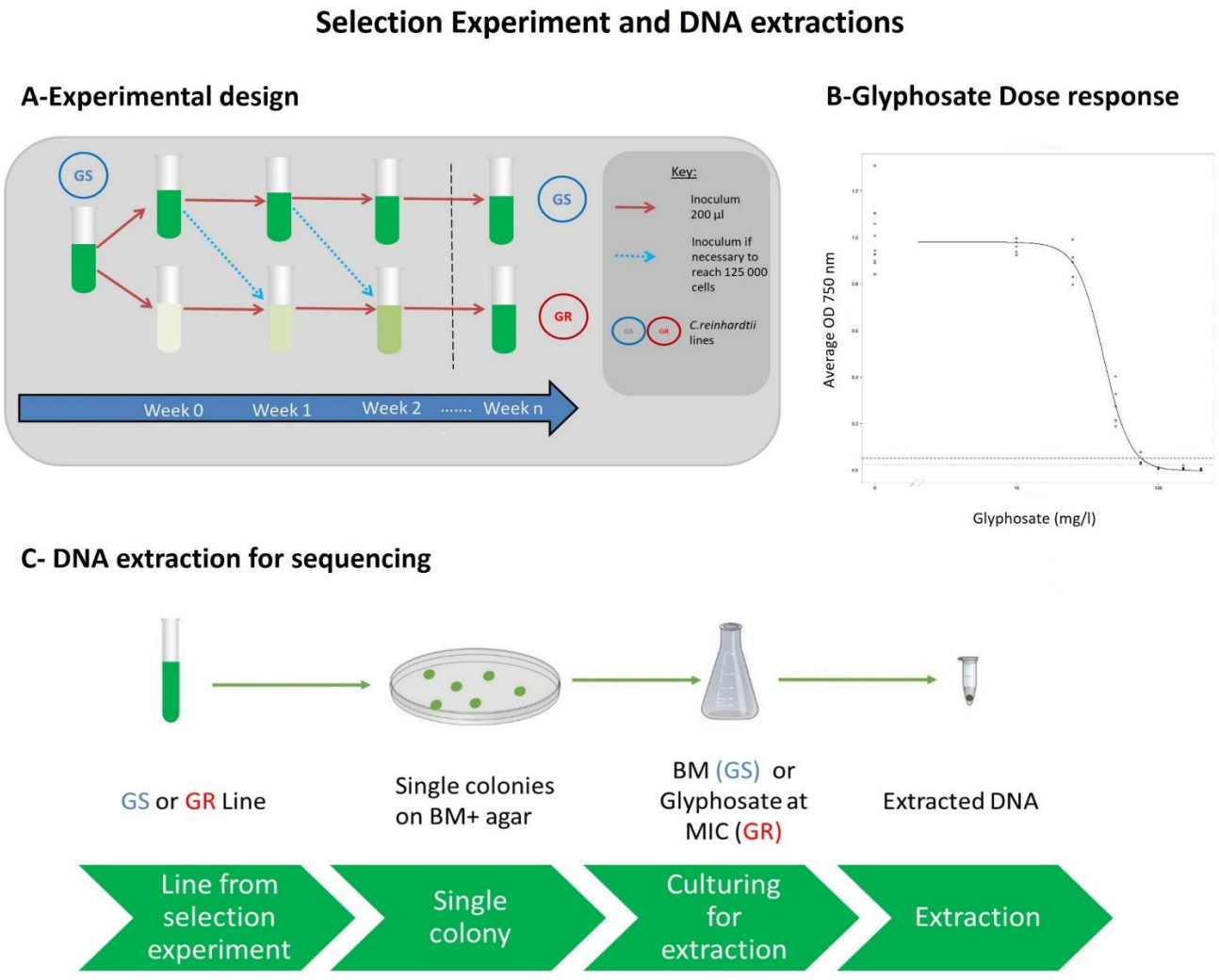

Figure 6: Glyphosate selection experiment design and DNA sampling approach for whole genome sequencing. (A) Selection experimental source-sink design (for detail, refer to Chapter 2): five distinct glyphosate sensitive lines of *C.. reinhardtii* CC-1690 strain were used to inoculate non-selective media (BM) and glyphosate selective media resulting in five Glyphosate sensitive (GS)/Glyphosate resistant (GR) replicate pairs. (B) Glyphosate dose response experiment: A dose response experiment had previously established glyphosate minimum inhibitory concentration (MIC = 100 mg. l^{-1}). (C) DNA sampling approach for whole genome sequencing of three GS/GR replicates: Six samples were plated on agar to isolate single colonies. A single colony per sample was then cultured in 200 ml of appropriate liquid media for four days, centrifuged to collect cells for DNA extraction adapted for *C. reinhardtii* (for detail refer to Chapter 2).

3.2.3 Library preparation and sequencing

Three GS/GR replicates were chosen for whole genome sequencing. The six lines were streaked to get single colonies to inoculate 200 ml of fresh BM (GS) or BM containing glyphosate at the MIC (GR) as described in Figure 6C. The six cultures were also checked for contamination on nutrient agar plates and were all axenic. Cells from all six lines were harvested for DNA extraction and DNA was extracted using an in-house DNA extraction protocol for Chlamydomonas (for details see Chapter 2 section 4.3 and/or Appendix 2). DNA integrity was assessed on a gel and purity determined on a nanodrop spectrophotometer. Quantity was determined by fluorometric quantification using the Qubit dsDNA BR Assay Kit. PCR-free library preparation and sequencing was performed on all six samples by the Earlham Institute using an Illumina HiSeq2500 sequencer which produced 250bp paired-end reads and average coverage of 100X. Sample names and replicates are as detailed below (GS = glyphosate-sensitive , GR = glyphosate-resistant):

Replicate	Sample name:
A	GSA
	GRA
B	GSB
	GRB
C	GSC
	GRC

3.2.4 Variant calling and genotyping

General steps, tools and their parameters for the variant calling and genotyping pipeline are described in detail in Chapter 2. Here are detailed the results specific to sequencing of the six samples from the SR selection experiment described in section 3.2.2 .

3.2.4.1 Read processing

Raw read quality was assessed (FastQC v0.11.9) and read pre-processing (Trimmomatic v0.39) was performed to exclude Illumina TruSeq2-PE adapter sequences. Read ends were then further trimmed for quality with a minimum window size of 5 and quality score of 20. Reads shorter than 50 bp after trimming were removed. Post trimming read assessment (FastQC v0.11.9) confirmed high overall paired-end read quality, with Phred scores > 30 across the read length. GC content matched the expected genome composition, and duplication levels were low. Across the six samples, 87.95 to 89.89 % of the reads were retained as high-quality reads for further analysis (Table 5). Trimmed sequencing data was deemed suitable for subsequent variant calling analysis.

Table 5: Number of reads for each of the six samples included in the SR sequencing experiment , with percentage of reads left after processing to retain only high-quality reads. Values were obtained from Trimmomatic (v0.39).

Samples	GSA	GRA	GSB	GRB	GSC	GRC
Number of raw reads (in millions)	33.5	29.7	25.6	27.1	28.4	29.0
Reads retained after trimming (%)	88.43	88.77	87.95	88.37	88.73	89.89

3.2.4.2 Reference genomes

A comparison of two *C. reinhardtii* reference genomes for use in alignment and variant calling was done initially as there were perceived pros and cons of using one or the other. Both reference genomes were assembled from laboratory strains CC-503 and CC-1690. Both strains are derived from the first isolate collected by Smith in 1945 and their genealogy is well known (Harris, 2008b; Pröschold et al., 2005). The first reference used, here referred to as Phytozome-CC503_V5.6, was the V5.0 assembly of the CC-503cm92mt+ strain and its v5.6 annotation available on JGI's Phytozome13 portal (Merchant et al., 2007) (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii). Phytozome-CC503_V5.6 is the latest assembly and annotation for this strain available on Phytozome to date. The second reference genome used, here referred to as O'Donnell-CC1690_V1.0, is a highly contiguous nanopore nuclear genome assembly of our laboratory strain CC-1690 published in 2020 (O'Donnell et al., 2020). This assembly contains five more complete benchmarking universal single-copy orthologs (BUSCOs) than the Phytozome-CC503_V5.6 reference (O'Donnell et al., 2020). Both genome reference sequences were modified to include chloroplast and mitochondrial genome sequences from CC-503cm92mt+ genome version 3.1 (archived on Phytozome https://phytozome-next.jgi.doe.gov/) to avoid misalignment of organelle-derived reads to other parts of the nuclear genome reference (Ness et al., 2015).

The assembly sizes of both reference genomes (Phytozome-CC503_V5.6 and O'Donnell-CC1690_v1.0) were similar (Table 6).

Table 6: Genome characteristics for the three references tested. Values were obtained from Quast (Galaxy Version 5.3.0+galaxy0 on galaxy EU).

Reference genome	Phytozome-CC503_V5.6	O'Donnell-CC1690_V1.0
No. of contigs	68	32
Largest contig(Mb)	9.731	9.805
Genome size(Mb)	111.401	111.412
N50(Mb)	7.784	6.886
L50	7	7
N90(Mb)	3.827	4.016
L90	15	15
GC (%)	64.02	64.07

To compare performance of both reference genomes, I aligned reads from the SR experiment to both references with BWA-MEM (v0.7.17), obtained alignment quality metrics from Qualimap (v2.2.1) and compared the results (Table 7). All read alignment quality metrics were preferable using the O'Donnell-CC1690_v1.0. While most improvements were only minor, there were clear differences in the standard deviation of the coverage. The most likely explanation for such differences are known structural genome differences between the two strains (Flowers et al., 2015; Gallaher et al., 2015).

Table 7: Alignment metrics averaged across the six samples of the Single Rate (SR) selection experiment for each of the two reference genomes tested. Values were obtained from Qualimap (v2.2.1).

Reference genome	Phytozome-CC503_V5.6	O'Donnell-CC1690_V1.0
Mean read length	230.50	231.19
Read mapped (%)	99.32	99.87
Properly paired reads (%)	98.99	99.81
Mean coverage	105.24	106.29
Standard deviation coverage	227.55	124.56
Mean mapping quality	52.27	53.27

Initial variant calling results were further filtered (see Chapter 2 for detailed information on filters) to retain high quality variants that may be linked to glyphosate resistance (Table 8). After this more stringent filtering the list of variants potentially linked to glyphosate resistance called using the Phytozome-CC503_V5.6 reference still contained numerous false positives: the alternate allele frequency in some GS samples were too high to confidently support a glyphosate-resistant variant and were deemed to be false positives. In contrast, the list of variants susceptible to support glyphosate resistance was very much improved as such false positive variants were not detected when using the O'Donnell-CC1690_v1.0 reference (Table 8). For these reasons, the variant list produced using the O'Donnell-CC1690_v1.0 reference assembly was used in downstream analyses of candidate variants.To facilitate further analysis of variant lists using annotation-based methods, the O'Donnell-CC1690_v1.0 genome was annotated by Rothamsted Genomics Service, transfering the Phytozome-CC503_V5.6 to the O'Donnell-CC1690_v1.0 assembly using Liftoff v1.6.3 (Shumate and Salzberg, 2021).

Table 8: Number of variants called in each replicate GR/GS pair at different false positive filtering steps for each of the two reference genomes tested. Phytozome V5.6 reference genome has been assembled from CC-503 strain and contains additional chloroplast and mitochondrial genome sequences from ChleV3.1 reference genome also available on Phytozome (https://phytozome-next.jgi.doe.gov/) . O'Donnell CC1690_V1.0 reference genome has been assembled on CC-1690 strain and is available in NCBI (https://www.ncbi.nlm.nih.gov/nuccore/JABWPN000000000) and contains additional chloroplast and mitochondrial genome sequences from ChleV3.1 reference genome .

Reference genome	Phytozome-CC503_V5.6	O'Donnell-CC1690_V1.0
1-GS and GR replicate pair are different		
Rep. A	2550	1642
Rep. B	2392	1715
Rep. C	2223	1575
2-GR has non-reference call		
Rep. A	1500	1229
Rep. B	1719	1295
Rep. C	1579	1159
3-GR call is not in any GS sample		
Rep. A	737	538
Rep. B	767	490
Rep. C	749	492
4-Alternate allele frequency in GR is >0.85		
Rep. A	182	94
Rep. B	166	97
Rep. C	177	95

3.3 Results

3.3.1 Generating glyphosate sensitive (GS) and glyphosate resistant (GR) lines for genome analysis.

Throughout the experiment, all GS lines remained above the threshold value (a minimum cell density of 625,000 cells.ml^{-1}), confirming that growth conditions were satisfactory (Figure 7 A). In the first three weeks of selection, the biomass of all GR lines remained under the threshold. After 15 weeks of selection, all GR lines showed and maintained a glyphosate resistance phenotype (Figure 7 B). Sequencing costs prohibited analysis of all five GS/GR sample pairs. GR lines derived from GSA, GSB and GSC exhibited contrasting phenotypic trajectories under selection by evolving resistance at different times during the selection experiment (GRC:3 weeks, GRA:5 weeks, GRB:10 weeks). For this reason, these lines were considered interesting for further study and were thus cultured for DNA extraction and whole-genome resequencing.

3.3.2 Analysis of the EPSPS-encoding gene

As a known host of target-site resistance mutations in other organisms, the EPSPS-encoding gene (Schönbrunn et al., 2001) was identified in the Chlamydomonas genome and analysed for the presence of mutations. To identify putative variants located within or nearby to the *EPSPS* gene, any variant calls annotated by SnpEff as putatively affecting the relevant gene model (*Cre03.g181300)* gene were extracted. Only one variant site was identified. An insertion of four guanines within a predicted intron (chromosome 3, position 5093521) was called as a variant against the reference genome for all six samples. Absence of any difference between the GS and GR genotypes at this locus suggests that this mutation does not confer glyphosate resistance.

To look for evidence of *EPSPS* copy number variation between GS and GR samples, counts of read alignments for this gene were extracted from the bam files, normalised to account for differences in the amount of sequencing data available, and ratios compared for the three GS/GR pairs. Alignment counts ratios (GS/GR) were 0.96, 0.95 and 1.01 for replicates A, B and C, respectively, indicating that there was no evidence to support *EPSPS* copy number change in this experiment.

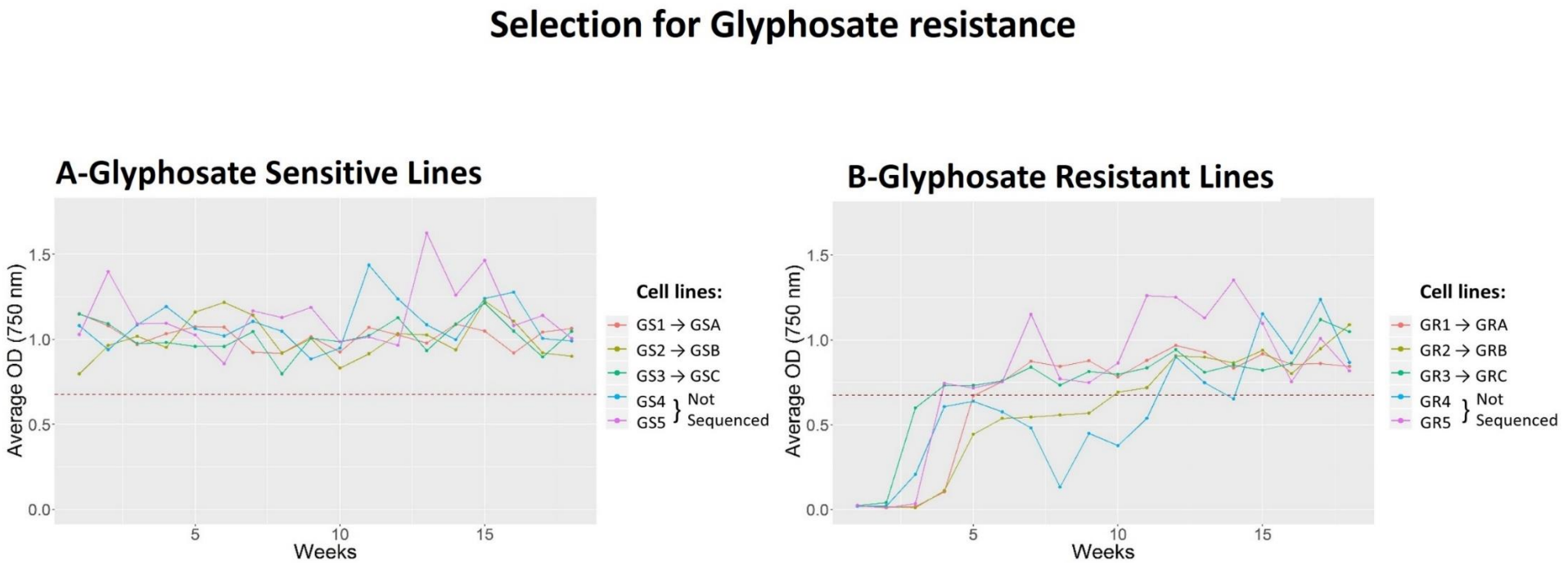

Figure 7: Selection for glyphosate resistance generating GS and GR lines for genome analysis. (A)Average optical density after 7 days on GS lines over 18 weeks. (B)Average optical density after 7 days on GR lines over 18 weeks. Horizontal dashed line represents the threshold OD_{750} value above which lines are considered resistant to glyphosate: cell density is high enough to ensure that a minimum of 125,000.

3.3.3 Genome-wide analysis of variants and annotations

The filtered list of putative variants supporting glyphosate resistance comprised 94, 97 and 95 variants for replicate pairs A, B and C, respectively (see Appendix 3). The average variant calling quality and genotype qualities (Table 9) are high and largely consistent across the replicates. On average the calculated alternate allele frequencies (AAF) are as expected for variants supporting glyphosate resistance: variant alleles frequencies were close to 1 in the GR samples on average (as expected from the > 0.85 allele frequency filter used) and very low in the GS samples, on average.

Analysis of individual variant calls that passed all filters used indicated that some calls had noticeable different characteristics, when compared against the entire filtered variant data set.

Table 9: Average alternative allele frequencies (AAF) and average genotype qualities (GQ) for glyphosate resistant and glyphosate sensitive samples in each of the three replicates of the SR experiment.

Replicate	Samples	Average AAF	Average GQ
A	GR	0.976	148.82
	GS	0.003	152.03
B	GR	0.982	147.03
	GS	0.019	149.73
C	GR	0.981	148.43
	GS	0.025	147.60

Few variants of lower confidence presented a higher alternate allele frequency (GS AAF > 0.1) in the GS sample (Figure 8 A) and few variants presented a lower genotype quality (GR GQ < 130) in the GR samples (Figure 8 B). Although these calls could potentially be removed in next iterative filtering process, they were left in the final list of variants for the time being.

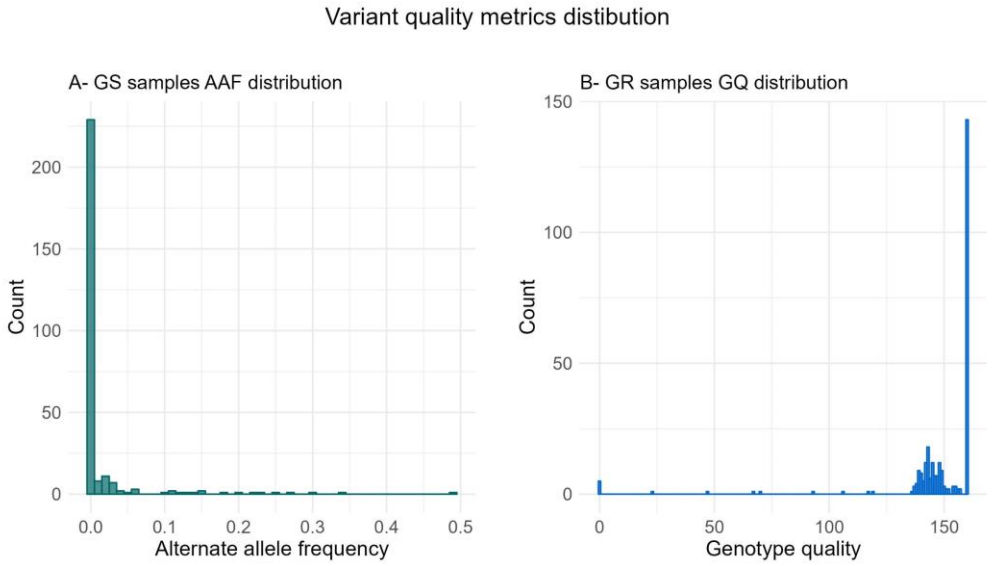

Figure 8: Distribution of variant quality metrics in list of variants. (A) Distribution of alternate allele frequency (AAF) in the glyphosate sensitive (GS) samples.(B) Distribution of genotype quality in the glyphosate resistant (GR) samples.

SnpEff variants annotation provided predicted annotations, impact scores and gene names for the list of variants (Appendix 4).

3.3.4 Variant description

SnpEff (v4.3+Tgalaxy2) was used to categorise and annotate the final variant list based on the predicted effects of the genetic variants.

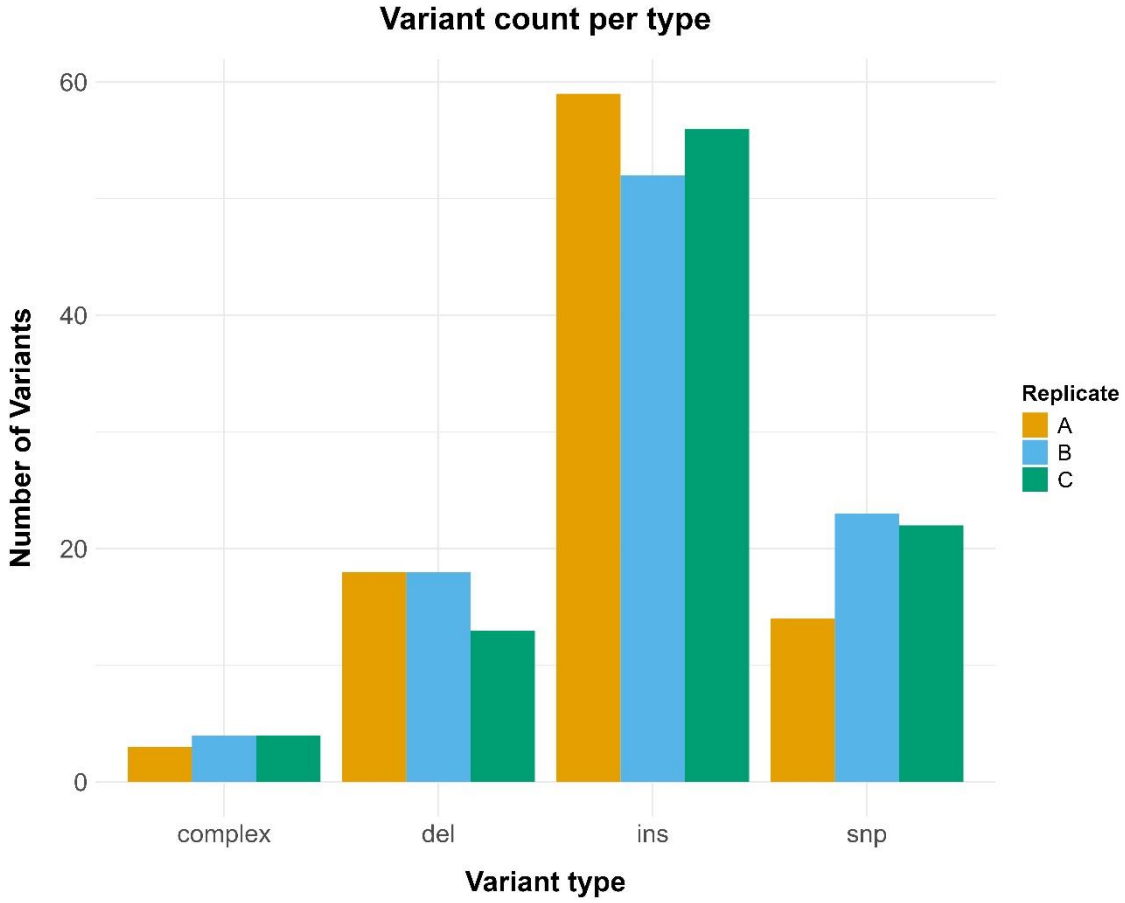

Figure 9: Variant type in each replicate from the SR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2). In snpEff, complex variants types are variants combining multiple types of change and do not fall into the simple categories of deletions (del), insertions (ins) and single nucleotide polymorphism (snp).

Single nucleotide polymorphisms (SNP), insertions, deletions and complex variants were found in all three replicates with proportions of each relatively stable across the different replicates (Figure 9). For each replicate, the predominant variant class was insertions.

When considering the positional annotation of variants relative to genes, the majority of variants were located in non-coding regions upstream, downstream of genes or within introns. No variants were annotated as located in a coding region (Figure 10).

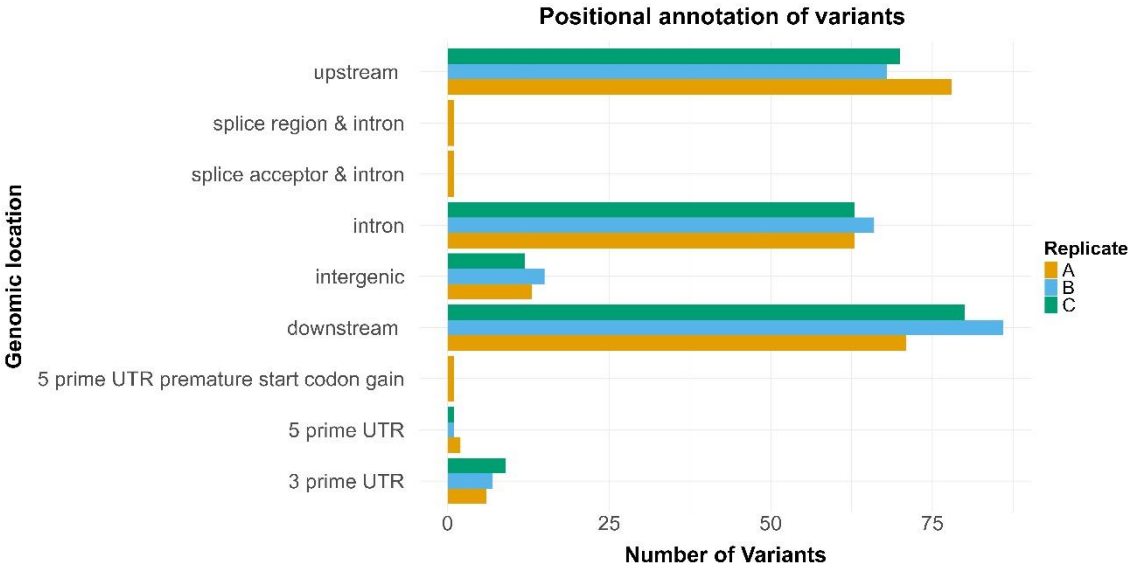

Figure 10: Variant location relative to the predicted impacted gene from the SR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

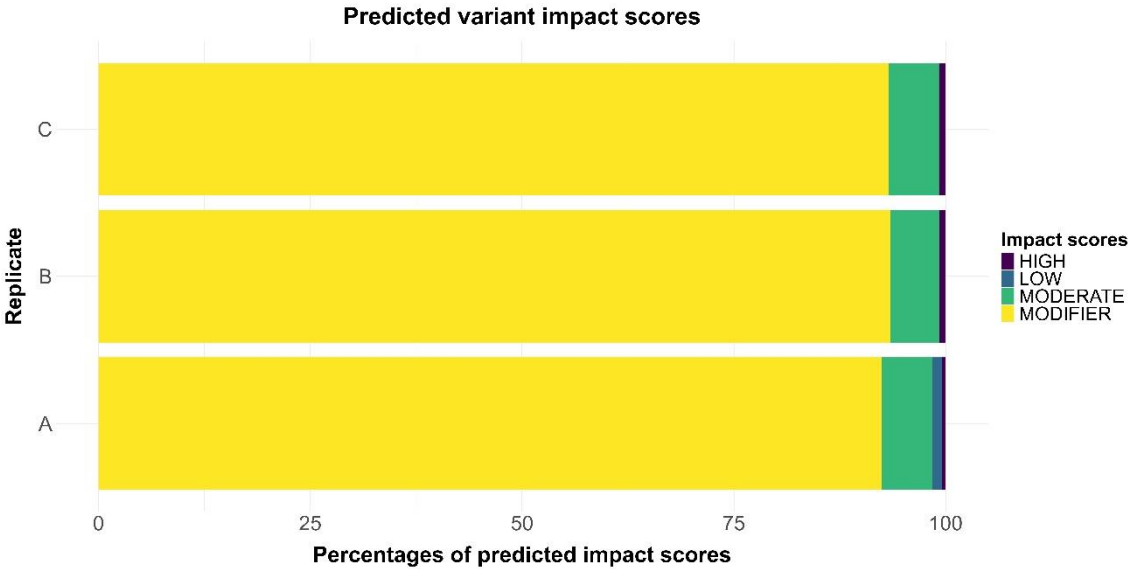

Figure 11: Predicted variant impact score from the SR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

The predicted effect of variants on the gene function is classified in four categories (in order of predicted impact magnitude: high, moderate, low and modifier) which represents a starting point to investigate their potential effect on the genotype. Few variants with predicted high impact scores were identified (Figure 11). Two genes, Cre01.g026350 and Cre05.g245150 that are predicted to be affected by high impact variants in replicates B and C have been assigned functional annotations while there is limited functional annotation available for the impacted gene (Cre08.g37495) in replicate A (Table 10).

Table 10: Loci with predicted high impact score variants from the SR selection experiment. Annotation with SnpEff (v4. 3+T.galaxy2). Gene name associated PlantFAMS information was extracted from Phytozome (https://phytozome-next.jgi.doe.gov/)

Chromosome	Position	Replicate	Annotation	Gene	Associated PlantFAMS	
					Viridiplantae	Chlorophyte
01	3980468	B and C	frameshift variant	Cre01.g026350	Protein Kinase	PF07707 - BTB And C-terminal Kelch (BACK)
05	635242	B and C	frameshift variant	Cre05.g245150	FAD dependent oxidoreductase, putative, expressed	PF13450 - NAD(P)-binding Rossmann-like domain (NAD_binding_8)
08	3012012	A	splice acceptor & intron variant	Cre08.g374950	Uncharacterized conserved protein	

3.3.5 Genomic distribution of variants

For all three replicates, no variants were reported in the non-chromosomal scaffolds after filtering steps to retain only variants putatively associated with glyphosate resistance. Variants are largely distributed across the genome (Figure 12). The distributions of variants of replicate sample B and C appear more similar. Thirteen genomic regions contain variants from the list that are common to all three replicates. They are potentially supporting glyphosate resistance or are located in problematic regions of the genome that generate false positives.

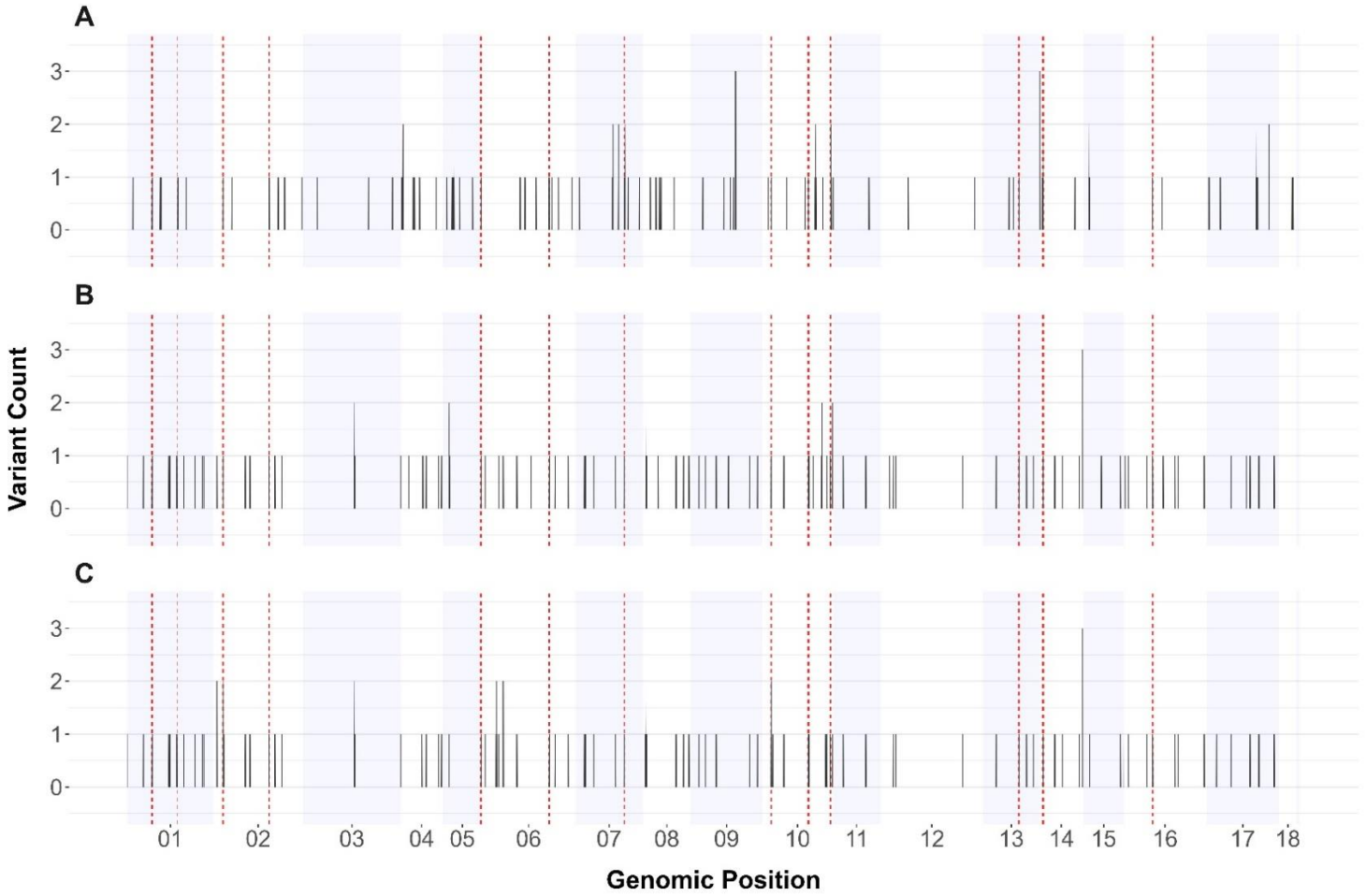

Figure 12: Variant distribution in each of the three replicates of the SR selection experiment. Variants were counted in overlapping bins (100 Kb sliding window with a 10 Kb step) along the genome. Shaded areas delimit chromosome boundaries. Vertical black lines represent bins with variants and their thickness reflects the number of consecutive bins in a given genomic region. Vertical red dashed lines represent regions with bins containing variants in all three replicates. Their thickness reflects the number of common bins in the genomic region.

3.3.6 Commonalities of variants between replicates

SnpEff annotation of the variant lists based on potential impact returned 475 gene models that may be affected by variants potentially linked to glyphosate resistance (Figure 13). These highlighted gene names were used as a basis for comparisons between replicates. Seven of these genes were common to all three replicates. Replicates B and C share a total of 193 of genes potentially linked to glyphosate resistance. The largest proportion of genes potentially affected (almost half of the total number identified across the experiment) were exclusive to replicate A variants.

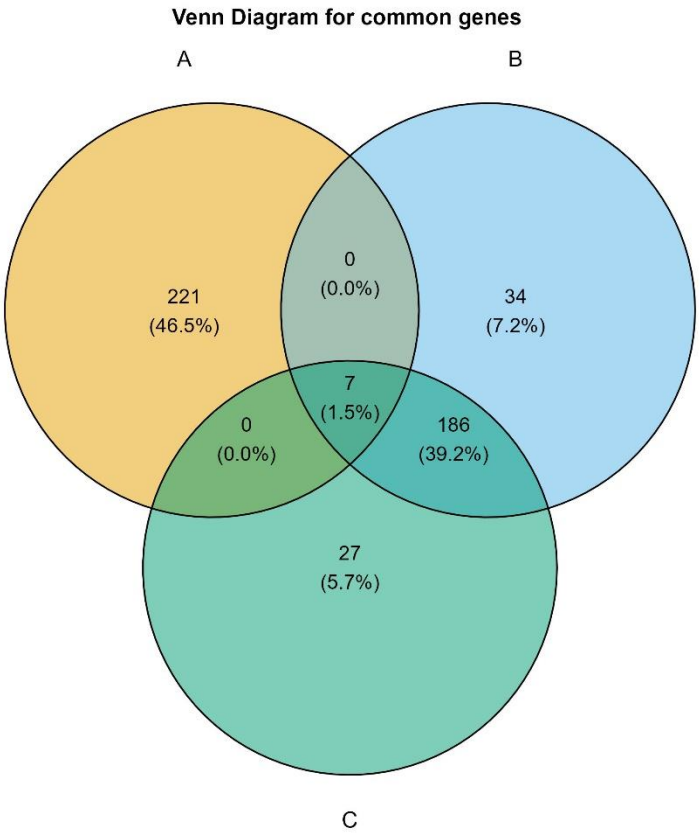

Figure 13: Number of gene names associated to variants susceptible to support glyphosate resistance in the SR selection experiment. Gene names from each replicate are colour coded (yellow: replicate A, blue: replicate B and green: C). Annotation via SnpEff (v4.3+T.galaxy2).

Table 11: Gene associated to variants susceptible to support glyphosate resistance common to the three replicates from the SR experiment. Annotation via SnpEff (v4.3+T.galaxy2) Description and Associated PlantFAMS information from Phytozome (https://phytozome-next.jgi.doe.gov/).

Gene	Description	Associated PlantFAMS	
		Viridiplantae	Chlorophyte
Cre01.g013800 (TCY1)	Tocopherol cyclase	zinc knuckle (CCHC-type) family protein	PTHR23002 - ZINC FINGER CCHC DOMAIN CONTAINING PROTEIN
Cre01.g013801	Tocopherol cyclase	tocopherol cyclase, chloroplast / vitamin E deficient 1 (VTE1) / sucrose export defective 1 (SXD1)	5.5.1.24 - Tocopherol cyclase
Cre05.g233702-CHR_END		PTHR11101 - PHOSPHATE TRANSPORTER	PF02689//PF05970 - Helicase (Herpes_Helicase) // PIF1-like helicase (PIF1)
Cre10.g450650		PF00168 - C2 domain (C2)	PF00168 - C2 domain (C2)
Cre10.g450700 (CSB39)	Probable transposon-derived protein of Chlamydomonas-Specific family B	PTHR15535 - TRANSMEMBRANE PROTEIN 2-RELATED	PTHR15535 - TRANSMEMBRANE PROTEIN 2-RELATED
Cre14.g610663		3'-5' exonuclease	3'-5' exonuclease
Cre14.g610700 (PHC72)	Pherophorin-chlamydomonas homolog 72	123394846	124076789

3.4 Discussion

The Evolve and Resequence (E&R) method is a powerful approach in evolutionary biology that combines experimental evolution with whole-genome sequencing to identify genetic changes associated with specific traits (Long et al., 2015). This approach allows researchers to observe evolution in real-time and pinpoint the genetic adaptations linked to the observed phenotype. E&R genomics studies are complementary to transcriptomics and other omics studies because they attempt to pinpoint the genetic changes, which might be in a distant locus, underlying the changes in gene expression and physiology observed with RNA-seq data.

In the context of glyphosate resistance, while E&R has been instrumental in studying resistance mechanisms in various organisms, its direct application to glyphosate resistance is not yet documented. Most studies on glyphosate resistance in plants have focused on mechanisms such as target-site mutations (Alcántara-de la Cruz et al., 2016; Morran et al., 2018; Ng et al., 2003; Perotti et al., 2019; Takano et al., 2020), and enhanced metabolic degradation (W. Deng et al., 2022; Pan et al., 2019; F. Y. Zhouet al., 2023). *C. reinhardtii*, a unicellular green alga, has been instrumental in studying glyphosate resistance mechanisms at the phenotypic level (Hansson et al., 2024; Lagator et al., 2012, 2013, 2014a; Melero-Jiménez et al., 2021; Vogwill et al., 2012). To the best of my knowledge, this study is the first to investigate links between evolution of a glyphosate resistant phenotype with concomitant changes in genotype using E&R.

Using experimental evolution, replicated pairs of glyphosate resistant (GR) and glyphosate sensitive (GS) lines were generated for genome analysis. Their sequences were compared to link the observed glyphosate resistant phenotypes to genomic changes using variant calling analysis and a filtering pipeline to only retain reliable variants putatively linked to glyphosate resistance.

Alignment and variant calling (VC) with the nuclear genome assembly for *C. reinhardtii* (O'Donnell-CC1690_v1.0), allowed development of an improved filtering pipeline to generate the final list of variants putatively linked to GR phenotype. After encountering difficulties using the Phytozome-CC503_V5.6 reference genome, the recent release of the highly contiguous O'Donnell-CC1690_v1.0 assembly (based on the sequencing of the same strain that was used in our selection experiment) was instrumental in the obtention of a final list of high-quality variants putatively linked to glyphosate resistance, reducing the list of potential variant by approximately 55% and removing false-positive calls. As costs diminish and methods improve, generating bespoke assemblies for one's laboratory strain might even improve resolution for this type of studies. Indeed, generating a genome assembly for a laboratory strain has been shown to significantly enhance variant calling analysis by providing a more accurate and specific reference genome. Bespoke high-quality assemblies for a strain provides improved reference genome specificity (Barbitoff et al., 2021), reduce mapping bias (Deng et al., 2021), and enhanced resolution of complex genomic regions (Li et al., 2023). These studies collectively demonstrate that generating a high-quality genome assembly for a laboratory strain enhances the accuracy, sensitivity, and resolution of variant calling analyses by providing a more precise reference, reducing mapping biases, and enabling better detection of strain-specific variants. The VC analysis resulting list was further filtered to only retain high quality variants potentially linked to glyphosate resistance. To produce this final list of variants involved filtering based on assumptions as detailed in Chapter 2. Choices were made to balance avoiding loss of true

variants (false negatives) and to minimise noise (falsepositives). Applying such filters may influence the results but is necessary to produce the final variant list for downstream analysis. While the final list of variants generated with the methods presented in this chapter contained a large number of high-confidence variants, it is likely to still contain false positives (FPs). Insertions represented the majority of variant types which contradict previous finding by Ness et al. (2015) who found indels to account for 36% of mutations. However, there is evidence that salt stress can lead to an increased rate of indel mutations in *C. reinhardtii* (Hasan et al., 2022). Therefore, while some insertions in the final list of variants may still be FPs, the high insertion frequency might also be the result of glyphosate exposure induced stress. To help resolve this and remove further potential FPs, refining the variant calling analysis by adjusting parameters for indel calls could be explored in future works. Additionally, the presence of microsatellites in the final list of variants (Appendix 3) also poses questions about the presence of FPs in the final list of variants. Microsatellites are short repetitive DNA sequences with high mutation rates due to DNA polymerase slippage during replication (Kelkar et al., 2010). These characteristics can lead to increased FP variant calls, complicating genomic analyses. However, there is some evidence that microsatellites can be involved in adaptation (Haasl & Payseur, 2013; K. Zhou et al., 2014), precluding a simple removal of these sites from consideration here. There was no evidence for or against keeping these potential FPs so I choose to be conservative and retain these variants as variants of lower confidence. Bearing in mind the limitations of the methods employed, and the variable confidence associated with some of the variants in the final list, I started investigating their putative links with the observed glyphosate resistant phenotype.

Glyphosate resistance in higher plant is supported by target-site resistance (TSR) and non-target-site resistance (NTSR) mechanisms. Glyphosate TSR mechanisms described in higher plants are EPSPS-encoding gene mutations, increased gene copy number, increased transcription or EPSPS increased activity (Galeano et al., 2016; Jander et al., 2003; Koo et al., 2018; Molin et al., 2017; Zhang et al., 2015). Glyphosate NTSR mechanisms described to date are reduced uptake, reduced translocation, enhanced metabolic degradation and, detoxification (Deng et al., 2022; Michitte et al., 2007; Pan et al., 2019; Vila-Aiub et al., 2011). There is no evidence to date that heritable epigenetic modifications are implicated in glyphosate resistance in weeds (Sen et al., 2022).

Preliminary investigation of this variant list has been conducted focusing on small variants using Freebayes, a widely used haplotype-based variant detector designed to identify small polymorphisms (single-nucleotide polymorphisms (SNPs), insertions and deletions (indels), multi-nucleotide polymorphisms (MNPs), and complex events (Garrison and Marth, 2012)).

There was no evidence for variants supporting a GR phenotype affecting the EPSPS-encoding gene in the final list of variants putatively linked to GR phenotype. As a result, the presence of *EPSPS* target-site mutations widely reported in higher plants, does not appear to be the source of the evolved glyphosate resistance phenotype in the current experiment. Also, a preliminary informal investigation of read depth around the *EPSPS* loci did not suggest the presence of *EPSPS* copy number variation (CNV) in the GR samples. These results suggest that if glyphosate TSR is involved, it is likely through mutation(s) in a distant genomic region(s) affecting *EPSPS* expression. TSR and NTSR mechanisms can co-exist in single individuals and populations, therefore mutations underpinning glyphosate NTSR mechanisms are likely to be present in the final list of variants. This result contrast with the assumption made by Vogwil et al. (2012)

based on the observed phenotype of glyphosate resistant *C. reinhardtii* populations exhibiting low fitness cost and absence of cross-resistance to other herbicides, which they postulated would be more likely to be indicative of TSR mechanisms.

The next logical step with the available dataset to rule out TSR mechanisms implication in the observed GR phenotype in the SR experiment would be to conduct variant calling to detect structural variants (SVs, i.e. larger genomic alterations) and search for evidence of SVs around *EPSPS*. However, SVs analysis requires specialized tools for identification and accurate characterisation. This analysis was not done here due to time constraints, but could be achieved in future work, using the Genome Analysis Tool Kit (GATK) which includes specialised tools designed for SV detection such as StructuralVariationDiscoveryPipelineSpark and GATK-SV pipeline (Caetano-Anolles, 2024). Similarly, a future study reproducing conditions in the SR selection experiment could investigate differential expression of the EPSPS-encoding gene between GR and GS lines, to further examine if or how upregulation of this gene plays a role in *C reinhardtii* glyphosate resistance.

NTSR resistance mechanisms are more challenging to characterise due to their diversity and polygenic nature (Délye and Christophe, 2013). In this thesis chapter, the final variant list has been explored using a range of simple diagnostic tools and comparison of patterns between replicates, to examine the potential role of individual variants in supporting a GR NTSR mechanism.

SnpEff annotations were used as a first point of investigation. In the final list, no variant was annotated by SnpEff as being located in a coding region, suggesting that variants linked to the observed glyphosate resistant phenotype may be found in regions impacting gene expression (such as UTRs, enhancers, silencers, transcription factor binding sites, promoter regions etc.). This is consistent with the fact that only three variants were classified as having a predicted 'HIGH' impact on gene function, likely through disrupting gene's reading frame (frameshift variants). These results suggest that the observed putative variants linked to GR phenotype might not lead to amino acid changes in encoded proteins. NTSR can be a polygenic trait, meaning it results from the combined effect of multiple genes, each potentially contributing a small effect. Mutation(s) underpinning an NTSR mechanism may be located in a metabolic encoding gene coding region and thus potentially impacting the protein sequence, or mutations (s) are located in a distant loci and potentially impacting metabolic encoding gene expression. These genes are often part of existing stress-response pathways within plants implying that variant with a large effect on encoded protein is more likely to be lethal and not contribute to a resistant phenotype. Under herbicide selection pressure, alleles conferring minor resistance can accumulate over generations, leading to a NTSR resistant phenotype. This gradual accumulation allows weed populations to adapt to herbicides even without direct mutations (of larger effect but without alteration to the enzyme function) in the target-site (Délye et al., 2013; Loubet et al., 2023). This polygenic nature makes NTSR more complex as it doesn't rely on single, easily identifiable mutations.

Major limitations come with using SnpEff annotations to interpret variant data. First, like any algorithm these are only predictive, and although potentially informative, further functional validation to assert the role of a variant in the observed phenotype will often be required. In *C. reinhardtii,* methods such as CRISPR/Cas9 gene editing (Ghribi et al., 2020), RNA interference (Cerutti et al., 2011), and gene overexpression (Hema et al., 2007) can be used for functional

validation of gene function. Secondly, SnpEff predictions rely on the quality of the available reference genome assembly and annotations.

Taking a different approach to variant investigation, genomic variant distribution was investigated for the presence of variant clusters, which could signal obvious “glyphosate resistance hotspot/s”. A similar approach was applied to explore distribution of DEGs and identify genomic regions linked to herbicide resistance in *Amaranthus tuberculatus* (Giacomini et al., 2020). There were no obvious clusters of variants, indicating no glyphosate hot spots in the genome of any of the three replicates studied. Consequently, the number of variants in the final list could not be narrowed down further (based on genomic position information) to fewer candidate genes to take forward for functional validation. However, from a final list containing 286 putative variants, I choose to use the list of genes impacted by variants with high predicted impact and genes impacted by variants common to all three replicates as a shorter list of nine genes to link with available functional annotations and their possible link with GR phenotype. Functional annotation available for the genes predicted to be impacted by these variants suggest that some variants may play a role in stress tolerance (Table 12). These candidate genes are closely linked to various stress tolerance mechanisms in plants and algae. They encode proteins involved in DNA repair, oxidative stress mitigation, cell signalling, extracellular matrix remodelling, and protective metabolite biosynthesis. These processes are essential for adaptation to environmental challenges such as heat (Nguyen et al.,2014, Vargas-Blanco and Shell, 2020, Ye et al.,2020), drought (Sun et al,.2021, Wang et al., 2020, Lim et al., 2020), salinity, and oxidative stress (Jain et al., 2020, de Carpentier et al., 2022). These genes highlight key molecular pathways that *C. reinhardtii* may be activating under glyphosate exposure. One particularly noteworthy gene is *Cre05.g245150,* classified as a putative FAD dependent oxidoreductase and predicted to be potentially affected by ’high' impact variants in replicates B and C (Table 10) by SNPeff. There is evidence suggesting that oxidoreductases (in fungi) and glyphosate oxidoreductase (GOX - in bacteria) play a role in glyphosate degradation (Firdous et al., 2020; Guo et al., 2021; Zhao et al., 2015). Although functional validation is necessary to ascertain the role of these variants in supporting the glyphosate resistant genotype, such results provide preliminary evidence that existing stress response pathways in *C. reinhardtii* may be implicated here in the observed evolutionary response to glyphosate. Furthermore, these results indicate that the method developed in this chapter could be successful in linking relevant variants to the observed glyphosate resistant phenotype provided further improvements are implemented.

Table 12: Subset of putative "Candidate genes" of the SR experiment with described gene function. Gene information from SnpEff (v4.3+T.galaxy2). Descriptions of putative gene product from various sources (Phytozome-CC503_V5.6 annotations or PlantFAMS) was retrieved from JGI's Phytozome13 Portal (https://phytozome-next.jgi.doe.gov/). Comments summarise potential links to stress tolerance from peer-reviewed sources.

Description	Description source	Gene	Comments
PIF1-like helicase	(PlantFAMS-Chlorophytes)	*Cre05.g233702*-CHR_END	PIF1-like helicase are multifunctional enzymes that play a role in maintaining genome integrity in eukaryotes (involved in DNA repair and replication stress response pathways) (Boulé and Zakian, 2006; Muellner and Schmidt, 2020). Replication stress (when the normal progression of DNA replication is impeded that can lead to DNA damage and genome instability).
3'-5' exonuclease	(PlantFAMS-Chlorophytes& Viridiplantae)	*Cre14.g610663*	In plants, 3'–5' exonucleases are enzymes that degrade RNA molecules from their 3' end, playing a crucial role in regulating gene expression and maintaining RNA homeostasis. There is evidence that 3'-5' exonucleases play a role in heat stress tolerance in *Arabidopsis* (Nguyen et al., 2014) and *E. coli* (Vargas-Blanco and Shell, 2020).
C2 domain	(PlantFAMS-Chlorophytes& Viridiplantae)	*Cre10.g450650*	C2 domain first identified in protein kinase, targeting protein to cell membranes resulting in signalling cascades in response to salt (Fu et al., 2019; Sun et al., 2021), drought (Sun et al., 2021) and heat (Ye et al., 2020) stress in plants.
FAD dependent oxidoreductase, putative, expressed	(PlantFAMS-Viridiplantae)	*Cre05.g245150*	FAD-dependent oxidoreductases (such as gluthathione reductases) are integral to plant stress tolerance, participating in various metabolic pathways that mitigate the adverse effects of environmental stresses. There is evidence of FAD-dependent oxidoreductase role to various abiotic stresses (Gill et al., 2013) such as high salinity, oxidative stress(Jain et al., 2020).
Pherophorin-chlamydomonas homolog 72	(Phytozome V5.6 annotation)	*Cre14.g610700* (PHC72)	Pherophorins are a family of ECM proteins found in green algae (Hallmann, 2006; von der Heyde and Hallmann, 2023). In *C. reinhardtii* they contribute to the formation of multicellular aggregates, a response to abiotic stresses such as high salinity or oxidative stress (de Carpentier et al., 2022).
Protein Kinase	(PlantFAMS-Viridiplantae)	*Cre01.g026350*	Protein kinase (PKs), by modulating the activity of specific proteins, orchestrate complex responses that enhance plant survival and adaptation under challenging environmental conditions. Some PKs are activated by abiotic stresses in plants (Majeed et al., 2023), others have been found to play central roles in drought, osmotic (Wang et al., 2020) and

			temperature stress response (Praat et al., 2021).
TMEM2- related	(PlantFAMS-Chlorophytes& Viridiplantae)	*Cre10.g450700* (CSB39)	TMEM2 (in plants homologs are named TMEM2-related proteins) is a trans membrane protein that degrades and regulates levels and function of hyaluronan (aka hyaluronic acid HA) (Yamamoto et al., 2017). HA is present in the extracellular matrix (ECM) and plays a crucial role in maintaining cell wall integrity, cell signalling facilitation and modulating responses to environmental stresses. There is evidence that HA plays a multifaceted role in stress tolerance by modulating inflammation (Petrey and de la Motte, 2014), protecting against oxidative damage and facilitating tissue remodelling (Berdiaki et al., 2023) in humans.
Tocopherol cyclase	(Phytozome V5.6 annotation) (Phytozome V5.6 annotation)	*Cre01.g013800* (TCY1) *Cre01.g013801*	Tocopherol cyclase is involved in vitamin E biosynthesis and play a pivotal role in plant cell protection from oxidative damage (Kanwischer et al., 2005). Evidence suggests its role in salt (Ouyang et al., 2011), light and temperature (Niu et al., 2022) stress tolerance. They are localised in plastoglobules (structures associated with thylakoid membranes involved in lipid metabolism and storage (Vidi et al., 2006)).

Another aim of this study was to investigate the repeatability of evolutionary outcomes under glyphosate selection in *C. reinhardtii,* both at the phenotypic and genomic level.

All lines under selection exhibited a glyphosate resistant phenotype by the end of the 18 weeks of the SR glyphosate selection experiment. Although the outcome of selection was identical across replicates, the three replicates chosen for sequencing presented quite different evolutionary dynamics. The different replicate lines differed considerably in the time taken to evolve resistance over the course of the selection experiment (replicate C: 3 weeks, A: 5 weeks, and B: 10 weeks). When comparing the genotypes linked to the observed GR phenotype, genomic changes appeared more similar for replicates B and C in terms of variant distribution along the genome and the list of genes predicted to be impacted by variants from the final list. Based on this data, multiple genotypes appear to support NTSR glyphosate resistance in *C. reinhardtii* in this experiment, and genomic similarity between putative GR variants is not necessarily related to similarity in time taken before resistance evolution. This result from a tightly controlled selection experiment using MEE is similar to what is described in the case of glyphosate NTSR resistance observed under much less controlled glyphosate exposure in agricultural systems (see Chapter 1 on GR mechanisms in weeds and microbes).

It is worth highlighting that any conclusion must take into account the limitations of the bioinformatics methods described above. For practical reasons, my analysis focused on a subset of the final variant list: specifically, variants classified as having high predicted impact (Table 10) , as well as those shared across all three replicates (Figure 13). In the latter case, the rationale was that genes harbouring mutations consistently across replicates are more likely to play a key role in glyphosate resistance. This targeted approach was particularly important given the novel variant-calling pipeline developed for this study; by concentrating on shared, high-impact variants, I was able to demonstrate that the resulting variant set included genes previously associated with stress tolerance in the literature, thereby supporting their potential relevance to the glyphosate resistance phenotype. However, evidence of independent evolutionary trajectories-particularly in replicate A-suggests that multiple, distinct genetic pathways may underlie resistance to glyphosate. Due to time constraints, a detailed investigation of these divergent responses was not undertaken but represents a clear avenue for future research. Additionally, the individual lines used during the selection experiment were not derived from a single cell and therefore clonality of their genome at the onset of selection can not be assumed even if they were the same strain.

The use of E&R here has allowed the investigation of small putative variants linked to an observed glyphosate resistant phenotype in *C. reinhardtii*. TSR mutations within the *EPSPS* coding region, common in GR in higher plants(De Carvalho et al., 2012; Collavo and Sattin, 2012; Deng et al., 2022; Galeano et al., 2016; González-Torralva et al., 2012; Han et al., 2016; Jasieniuk et al., 2008; Li et al., 2018; Ng et al., 2003; Perotti et al., 2019; Takano et al., 2020) were not observed, and nor were putative GR variants clustered into genomic hot-spots as-per Giacomini et al. (2020). Instead, several more general stress-related loci were implicated, with one consistent with previous studies of microbial glyphosate sensitivity (Firdous et al., 2020; Guo et al., 2021; Zhao et al., 2015), suggesting that NTSR glyphosate resistance mechanisms might be at play. Additionally, there was some indication of GR related variants being specific to individual replicate lines. These data suggest that although repeatability is observed among replicates at the phenotype level, convergence is not necessarily observed at the genomic level.

These conclusions must be considered in light of the limitation of the method developed and for which improvement have been suggested. Further improvement remains unattainable to date due to limited

available bioinformatics tools to address these questions. However, improvements to the selection experiment design could be implemented in future studies. First ensuring clonality of the genomes at the onset of selection to allow increase confidence in between-replicate comparisons. Another improvement of the selection experiment design would involve using a more realistic glyphosate selection scenario than application of a single and constant glyphosate dose.

4 Impact of rates of environmental deterioration on adaptation to glyphosate in *C. reinhardtii*

4.1 Introduction

Human-induced global environmental change is a current and major selective pressure affecting ecosystems and biodiversity globally, and the pace at which these changes induce biological evolution has accelerated over the past 50 years (Díaz et al., 2019). Understanding how populations adapt to environmental change is crucial to prevent biodiversity loss, for example due to climate change, as well as the emergence and spread of infectious diseases and pests (Gonzalez et al., 2013). More specifically, understanding cases when evolution results in population persistence in the face of environmental changes is important for both predicting future responses to environmental change, and furthering our fundamental understanding of evolution.

Environmental change moves a population from its current niche into new conditions where few individuals (if any) are able to survive and reproduce. In cases when population survival occurs, the evolutionary outcome is termed evolutionary rescue (ER) (Gonzalez et al., 2013). A considerable body of research has now been conducted on evolutionary rescue, including population adaptation to abrupt changes such as sudden exposure to herbicides (Kreiner et al., 2018), and population adaptation to gradual changes such as progressive climate warming (Schiffers et al., 2013). Common predictions currently are that in the first scenario, drastic initial reduction in population size is observed, such that beneficial mutations must be already present or rapidly arise via *de novo* mutations or immigration, to allow population survival. In the second scenario, reduction of population size is less noticeable and beneficial mutations are more likely to arise from standing genetic background and will be fixed sequentially following environmental changes (Gonzalez et al., 2013). It is considered that ER is more likely to occur in the second scenario (Bell and Gonzalez, 2011; Ferriere and Legendre, 2013).

Experiments to investigate the genetic mechanisms (such as changes in genome size, mutation rate, and gene expression and regulation) underlying ER and associated changes in organism fitness, are still needed (Gonzalez et al., 2013). Experimental evolution (EE) offers a powerful opportunity to study ER through well replicated experiments in tightly controlled environments. In particular, microbial experimental evolution (MEE) is a powerful approach to investigate the effect of environmental changes, allowing us to study evolution in action and uncover mechanisms of adaptation. Taking advantage of microbes' small size and short generation time and relatively simple genomes, large populations can be selected over several generations and used to study the effect of selection at both phenotypic and genotypic levels (Buckling et al., 2009; Elena and Lenski, 2003; Lenski, 2017a).

Herbicide applications are a human induced selective pressure on weed populations. Modern agriculture has relied mainly on herbicides for weed control, but in turn, this has driven the evolution of resistance to herbicides (Powles and Yu, 2010). Resistance is now a major threat to food security in the current

agricultural model (Oerke, 2006). As resistance is an evolutionary process, understanding the evolutionary ecology of adaptation to herbicide selection is crucial to future sustainable herbicidal use in agriculture (Neve et al., 2009, 2014). Herbicide resistance mechanisms are categorised in two groups: target-site resistant (TSR) and non-target-site resistance (NTSR) mechanisms (Gaines et al., 2020; Powles and Yu, 2010). TSR mechanisms are specialist to one herbicide mode action by limiting the impact on the target enzyme. This may be through mutation rendering the target enzyme insensitive (Gaines et al., 2020; Murphy and Tranel, 2019; Powles and Yu, 2010), or target gene over expression through increased transcription or gene copy number variation generating more target enzyme and thus increasing the herbicide dose required for effectiveness (Gaines et al., 2020). NTSR resistance mechanisms are not specialist to one herbicide mode of action (Gaines et al., 2020) and limit the number of herbicide molecules reaching the target enzyme (Gaines et al., 2020; Powles and Yu, 2010; Yuan et al., 2007). This may be through reduced uptake and translocation or enhanced herbicide degradation or detoxification (Gaines et al., 2020). Both TSR and NTSR resistance mechanisms can co-exist and confer higher resistance levels or resistance to multiple herbicides (Gaines et al., 2020; Powles and Yu, 2010).

The dynamics of herbicide adaptation will depend on the strength of selective pressure (Gressel, 2009; Powles and Yu, 2010) but studies have often focussed on the effect of sudden changes in herbicide selection, with exposure to high herbicide doses. The effect of lower herbicide doses rapidly selected for resistance in *Lolium rigidum* populations, which demonstrated cross-resistance to another herbicide mode of action and evidence towards a polygenic trait (Neve and Powles, 2005). It has been hypothesized that low pesticide doses causing stress to organisms could lead to enhanced mutation rates in survivors and consequently increase the number of resistant mutations in the population (Gressel, 2011). Gradual changes in dose that can occur with the buildup of compound released in the environment (Perron et al., 2008), however, remains to be further investigated.

Resistance management strategies often rely on the existence of a fitness cost (reduced plant fitness in the non-selective environment) being associated to evolved resistance (Purrington, 2000; Vila-Aiub et al., 2009). Fitness costs are a concept in evolutionary biology, whereby mutations or adaptations which are advantageous and therefore evolve under intense or rapid selection, may actually result in a lowering of the organisms' fitness in the absence of the selective pressure. One theoretical piece of evidence for this is that mutations endowing resistance are rare in the absence of herbicide, suggesting that they may represent a lower fitness than the wild-type genotype. Mutations in the target enzyme might interfere with important plant function or metabolism by reducing substrate affinity or catalytic capacity (Délye et al., 2005; Powles and Yu, 2010). Alternatively, NTSR mechanisms might come at an energetic cost by diverting resources that would otherwise have been allocated to growth and reproduction (Vila-Aiub et al., 2009). When compared to an herbicide sensitive individual in an herbicide free environment, reduced growth, reproductive success, or competitive ability in an herbicide resistant organism is considered evidence of a fitness cost (Vila-Aiub et al., 2009). The mechanisms underlying fitness costs are complex and can vary depending on the resistance mechanism, the genetic background, environmental and ecological conditions (Damalas and Koutroubas, 2024; Vila-Aiub et al., 2009), making their expression context-dependent (Comont et al., 2019b; Damalas and Koutroubas, 2024). Additionally, while fitness costs have been reported for evolution of herbicide resistance in a number of species and herbicidal modes of action (Han et al., 2017; Martin et al., 2017; Matzrafi et al., 2021; Menchari et al., 2008; Vila-Aiub et al., 2005; Yanniccari et al., 2016), there are equally a large number of cases where none has been detected (Giacomini et al., 2014; Keshtkar et al., 2017; Vila-Aiub et al., 2014; Yu et al., 2010)

Glyphosate is the most widely used herbicide in agriculture with 746,580 tonnes used globally, reported in 2014 (Antier et al., 2020). This extensive use has led to the evolution of glyphosate resistance in 60 weed species (Heap, 2024), posing a significant challenge to agriculture and polluting non-target ecosystems (Van Bruggen et al., 2018). Glyphosate inhibits 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), an enzyme of the shikimate pathway (Steinrucken and Amrhein, 1980) essential to the biosynthesis of aromatic amino acids in bacteria, algae, plants and fungi (Healy-Fried et al., 2007). In plants, glyphosate TSR mechanisms comprise either mutation of the *EPSPS* gene (Baerson et al., 2002), or *EPSPS* over expression due to either increased transcription (Zhang et al., 2015) or gene copy number variation (Molin et al., 2017). Glyphosate NTSR mechanisms reduce glyphosate dose reaching EPSPS through reduced uptake (Michitte et al., 2007), reduced translocation (Vila-Aiub et al., 2011), glyphosate detoxification (Pan et al., 2019) and degradation (Deng et al., 2022).

Although glyphosate resistance mechanisms are now relatively well characterised, our understanding of how fitness costs influence glyphosate resistance evolution remains limited: The degree of resistance and the associated costs depend on the species and the specific molecular mechanisms involved (Gaines et al., 2020; Sammons and Gaines, 2014). Additionally, factors such as trait dominance (Han et al., 2017), genetic background (Martin et al., 2017), life history stages (Osipitan and Dille, 2017), and environmental stressors like temperature (Ge et al., 2011) or competition (Pedersen et al., 2007) have substantial effects on the resulting phenotype. In some cases, resistance may be fitness-neutral (Vila-Aiub et al., 2014) or even provide a fitness advantage (Vogwill et al., 2012).

Some studies, suggest that changes to herbicide metabolism (NTSR) may indirectly influence plant fitness by diverting resources away from growth and reproduction and into defence mechanisms (resource-based allocation theory) (Vila-Aiub et al., 2019, 2009). A study demonstrated that fitness cost associated with NTSR ACCase resistance in black-grass can be mediated by ecological trade-offs (Comont et al., 2019b). In the case of glyphosate resistance, fitness costs associated with NTSR have not yet been extensively studied. Conversely, glyphosate TSR associated fitness costs have been comprehensively investigated and summarised in a review by Vila-Aiub, Yu and Powels (2019). In plants, *EPSPS* gene mutations are associated with mutation-specific biochemical trade-offs: in general mutations conferring lower levels of resistance to glyphosate (such as Pro106 substitutions and Thr-102Ser), do not come with significant fitness penalties. In contrast, mutations impairing EPSPS catalytic activity and conferring high levels of glyphosate resistance (such as Gly-101 substitutions, Thr-102-Ile, GAPS and TIPS mutations) are associated with higher fitness cost (Vila-Aiub et al., 2019). Second, *EPSPS* gene amplification and overexpression conveying glyphosate resistance in plants is not always associated with fitness costs, suggesting these may be genetic background-dependant or mediated by ecological factors (Vila-Aiub et al., 2019).

Glyphosate resistance and associated fitness costs have been studied with experimental evolution in *Chlamydomonas reinhardtii*, mostly using the CC-1690 laboratory strain exposed to a sudden increase in glyphosate doses to 1 MIC (Hansson et al., 2024; Lagator et al., 2012; Vogwill et al., 2012) and 0.5 MIC (Hansson et al., 2024). Selection experiment set ups were either serial weekly transfer in fresh media (Lagator et al., 2012; Vogwill et al., 2012) or continuous flow cultures (Hansson et al., 2024). Another study exposed a lake isolate to starting doses below MIC, increasing concentrations of glyphosate using a ratchet protocol and, successfully evolving resistance beyond 1 MIC (Melero-Jiménez et al., 2021). Every time heritable glyphosate resistance was observed, and authors concluded that it was acquired through adaptation. However, like in higher plants, glyphosate-associated fitness costs in *C. reinhardtii* were not consistently found across studies. A minor and a major glyphosate resistance associated-fitness cost was

reported in studies by Lagator et al (2012) and Melero-Jimenez et al (2021), while no fitness cost was reported in work by Hansson *et al* (2024) and Vogwill *et al* (2012). The mechanisms underlying glyphosate resistance and potentially associated fitness cost in *C. reinhardtii* remain unknown to date.

C. reinhardtii is a cosmopolitan Chlorophyceae with a world-wide natural distribution in both soil and freshwater ecosystems. It is also a model species for various areas of research due to its advantageous characteristics (Harris, 2008a). *C. reinhardtii* is susceptible to herbicides (Reboud, 2002), is already a model species to study herbicide resistance (Reboud et al., 2007) and evolutionary rescue (Lachapelle and Bell, 2012; Lachapelle et al., 2015b, 2017), and its genome has already been sequenced (Merchant et al., 2007). Recent work on evolution of herbicide resistance in *C. reinhardtii* aimed at understanding the effect of management practices (herbicide sequential application, cycling and mixture) and their underlying ecological and evolutionary theories (Lagator, 2012c; Lagator et al., 2012, 2013, 2014b; Vogwill et al., 2012). Most recently the effect of glyphosate dose on adaptation and associated fitness cost have been investigated in the *Chlamydomonas* 'CC 1690' strain (Hansson et al., 2024), as well as the effect of adaptation to increasing glyphosate doses in a lake isolate (Melero-Jiménez et al., 2021). The effect of increasing herbicide selective dose on evolution of resistance has not yet been investigated. A few microbial EE studies investigated the effect of continuously varying environments and suggested that it affected the dynamics and outcome of adaptation. The effect of decreasing phosphate concentrations was tested on *C. reinhardtii* populations (Collins and De Meaux, 2009), *Saccharomyces cerevisiae*, the model yeast, was exposed to increasing concentrations of heavy metals (Gorter et al., 2016) and *Pseudomonas* sp. and *E. coli* bacteria were exposed to increasing doses of antibiotics (Perron et al., 2006, 2008). Although most studies show comparable results in terms of evolutionary dynamics (i.e. sudden changes either delay or reduce the likelihood of adaptation), they do not always concur on the evolutionary outcomes (i.e. levels of fitness reached and associated costs). There is currently a gap in our general understanding of the effect of rates of change adaptation and associated fitness costs.

Using microbial experimental evolution (MEE), here I investigate the effects of the rate of environmental change on adaptation in the model organism *C. reinhardtii*, by exposing populations to various rates of glyphosate selection. Populations of *C. reinhardtii* were experimentally evolved in selective glyphosate environments. Different selective histories were created by varying the rates at which the glyphosate dose is increased to explore the effect on glyphosate resistance and associated fitness cost. This will test the hypothesis that:

(i) Rapid increases in glyphosate doses will reduce population size, increase variability and inconsistency of evolutionary dynamics and delay evolution of resistance.
(ii) Gradual glyphosate dose increase will allow for evolution of resistance to higher doses
(iii) Selective histories will affect fitness ancestral environment: Adaptation to gradual glyphosate dose increase will yield higher fitness cost.

4.2 Methods

4.2.1 Biological material and growth conditions

Axenic *C. reinhardtii* cultures maintained in our laboratory (CC-1690 wild-type mt+ [Sager 21 gr] strain) were kept in 20 ml of Bold's media (BM), as detailed in Chapter 2.

Population growth after 7 days is and estimate of fitness in the ancestral and selective environment and was inferred from estimates of population cell density (cell.ml^{-1}) by measuring optical density at 750 nm (OD_{750}) as detailed in Chapter 2. During the experiment, population cell density was estimated prior to each weekly transfer.

4.2.2 Glyphosate Stocks

A stock solution of 10 g l^{-1} of glyphosate was made directly in BM and used for the dose-response, selection, and resistance assay experiments. To ensure the stock solution did not degrade over time during the experiments, it was regularly tested with proton NMR and by exposing sensitive populations to the minimum inhibitory concentration (MIC). The six selective media were prepared weekly by adding 2 ml of 10X concentrated solutions into 18 ml of sterile BM.

4.2.3 Experimental design

4.2.3.1 Dose-response experiment

A glyphosate dose-response (DR) experiment was conducted to determine the glyphosate selective doses inducing a given percentage of inhibition in growth after seven days. To determine the selective doses to be used in the selection experiment, the glyphosate-sensitive wild-type *C. reinhardtii* was exposed to 10 glyphosate doses: 0, 25, 50, 75, 100, 125, 175, 200 and 225 mg l^{-1}. At each dose, six replicate populations (n=6) derived from the wild-type strain were assessed. Starting population size was 125,000 cells. Population density was estimated after seven days via measurement of OD750 as detailed in Chapter 2. The data is presented as a dose-response curve in Figure 14 , obtained by fitting a 3 parameter Weibull regression on the relationship between the glyphosate dose and the average OD750 (drm function of the drc package in R 3.5.0).

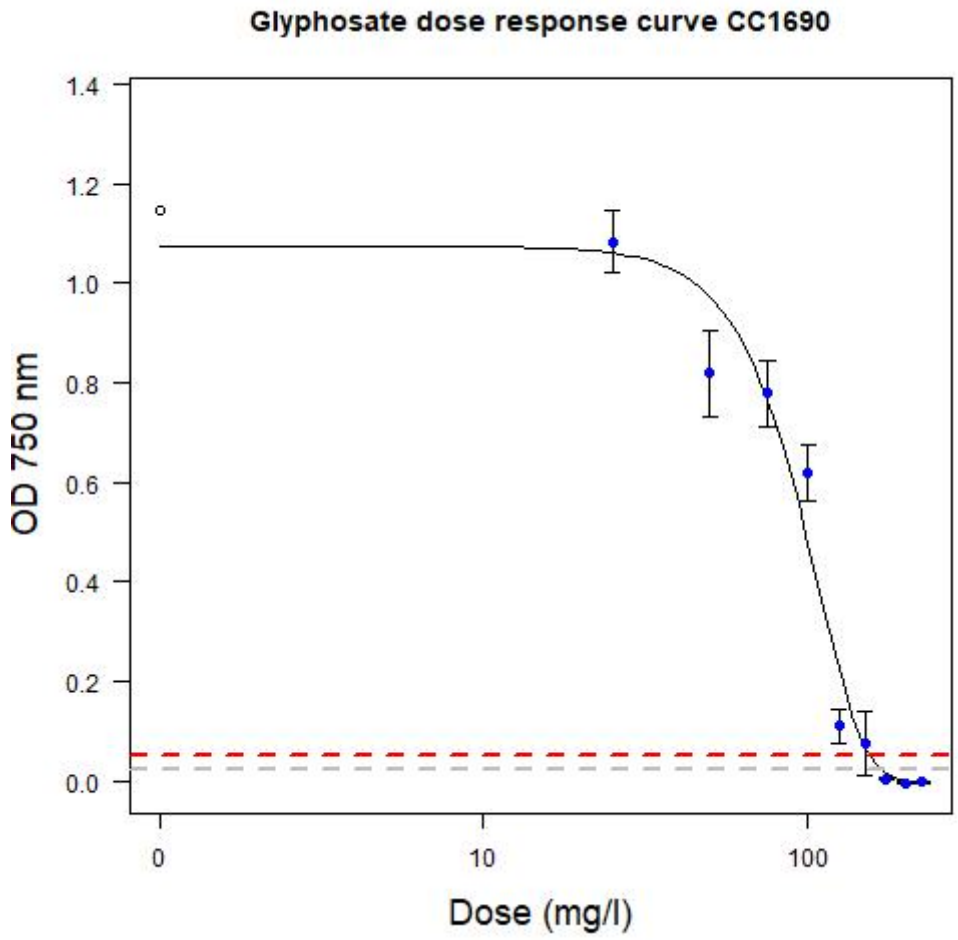

Figure 14: Dose response curve for 6 replicates of sensitive *C. reinhardtii* exposed to glyphosate for 7 days. The starting population size was 125,000 cells. Doses tested were 0, 25, 50, 75, 100, 125, 150, 175, 200 and 225 mg/l. Dots represent average OD_{750} at each dose with their standard errors. The solid black line is the three-parameter Weibull dose-response model returning the best fit (established with the mselect function from the drc package in R (3.5.0)).Grey dashed line represents detection limit. Red dashed lines represent threshold used to determine MIC (OD_{750} =0.05). .

A fitted three parameters Weibull dose-response model (drc package in R version 3.5.0) was used describe the relationship between the average OD_{750} and the glyphosate dose (Figure 14). To derive the minimum inhibitory concentration (MIC), which is the lowest glyphosate concentration inducing 100% inhibition of growth, along with doses leading to lower levels of inhibitions, the ED() functions from the drc package in R (3.5.0) was used. Doses above MIC can't be estimated from the ED() function, thus were calculated as a percentage of MIC. Table 13 presents the selective doses to be used for the selection experiment. They were either estimated or calculated as described above and then rounded up to the nearest 5mg/l.

Table 13: Glyphosate selective doses in relation to growth inhibition and MIC.Doses up to MIC were estimated using the ED() function from the drc package in R (3.5.0). All doses above MIC were calculated as a percentage of MIC. Selective doses to be applied in the selection experiment were estimated and rounded up to the nearest 5mg/l.

Glyphosate selective doses (mg/l)	Growth inhibition (%)	MIC
65	17	0.17
80	33	0.33
95	50	0.5
115	67	0.67
135	83	0.83
160	100	1
215	133	1.33
270	167	1.67
320	200	2

4.2.3.2 Selection experiment

A selection experiment (Figure 15) was run to test the effect of rates of change in glyphosate dose on the evolution of glyphosate resistance and the possibility to evolve resistance to higher selective pressures (above 1 MIC). After 12 weeks of selection, lines were used to investigate the effect of contrasting selective histories on levels of glyphosate resistance and associated fitness costs.

Eight source populations were kept in the ancestral environment (BM), to provide immigration into each of the eight replicates sink populations under selection as required (following a source-sink scenario, as detailed in Chapter 2), and served as controls (n=8). The selected populations were exposed to six glyphosate selection regimes with doses increasing up to either 1 MIC or 2MIC. For each of these endpoint doses, three rates of change were applied: quick (1 dose), intermediate (3 doses) and slow (6 doses).

During the selection experiment, both live cells and cells for DNA extractions were taken from each population at the start and every four weeks (four-time points). After the weekly transfer in fresh media, 100 µl of live cells were stored on BM acetate agar slopes and the remainder of the culture was harvested (centrifuged at 5500 g to discard media) and stored at -80°C. These samples were intended for use in a relaxed selection experiment. Unfortunately, due to ill health and the COVID-19 lockdowns, cultures stored on agar slopes did not survive.

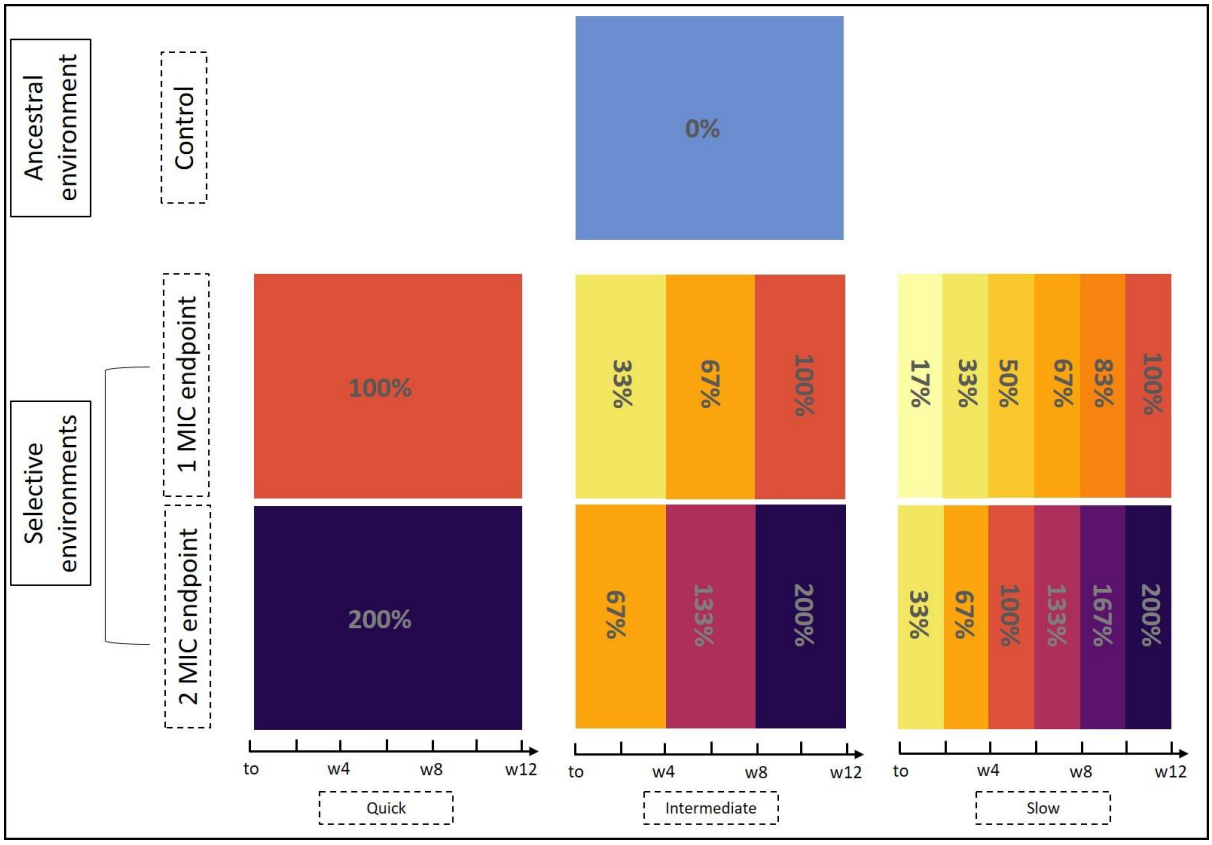

Figure 15: Selection experiment design.One isogenic sensitive population was used to inoculate all seven environments with eight replicates (n=8): ancestral environment (controls) and six glyphosate selective environments either with 1 MIC or 2 MIC as the highest dose. For each endpoint dose, three rates of change have been applied over 12 weeks. Quick change was imposed by selecting immediately with the endpoint dose, intermediate and slow change by gradual increase of doses towards the endpoint. Populations were transferred weekly into fresh media. DNA and live cells samples of the initial isogenic populations (t0) as well as the eight replicates in each of seven environments were collected at 3 additional timepoints (week 4,8 and 12) for further experiments (see Chapter 5).

4.2.3.3 Levels of resistance and fitness cost assays

After glyphosate selection (section 4.2.3.2), assays were established to ascertain the extent of glyphosate resistance and any associated fitness cost. First, to prevent glyphosate carry-over and ensure adaptation (rather than acclimation) accounts for the glyphosate resistance observed, all populations assayed were multiplied in ancestral (glyphosate-free) media for one week. Then two assays were conducted simultaneously to test the effect of selective histories on glyphosate resistance levels and fitness costs. An inoculum of 125,000 cells for each of the 32 populations that survived the selection experiment was transferred in glyphosate (1MIC) and ancestral media (BM). The assay included technical replication for 28 populations (four randomly selected population could not be duplicated due to lack of space) and was run twice (yielding a total of 240 observations).

To establish the effects of selection history on fitness in the presence or absence of the selective agent (glyphosate), population cell densities were used as a proxy for fitness: selected populations with a lower fitness are expected to duplicate more slowly, and therefore, have a lower cell density at the time of measurement. Cell densities were estimated from OD_{750} measurements during both exponential growth (after seven days) and at population's carrying capacity (after 14 days). These assays were replicated in a second run.

4.2.4 Data analysis

The effect of selective histories on population cell densities in the ancestral environment and in glyphosate at 1MIC were analysed separately, to investigate their effect on levels of glyphosate resistance and associated fitness costs. To achieve this, linear mixed models (LMM with lmer function of the lmerTest package using R3.5.0 software) were fitted.

Levels of glyphosate resistance was investigated after seven days (estimation of populations' growth rate). The expected death of all controls exposed to glyphosate at 1 MIC resulted in zero-inflated data distribution and they were excluded from analysis. Consequently, the model predictors were the six glyphosate selection regimes. The response variable was transformed as follows to improve homoscedasticity: $\log_{10}$(cell density +1). The 32 populations tested were considered as a random variable accounting for biological variation. Treatments (controls, quick, intermediate, and slow rates of change), replicate runs (2 levels) and blocks (2 levels) were considered fixed effects. When significant effects of selective histories were detected by the models, subsequent Least Significant Difference (LSD) multiple comparison tests (using predictmeans package in R 3.5.0 software) were run to test the differences between treatments.

Fitness in the ancestral media was investigated separately at both seven days (estimation of populations' growth rate) and 14 days (estimation of populations' carrying capacity). The predictors of the model were all the treatments (control and six glyphosate selection regimes). The response variable was the cell density following a log10 transformation to improve homoscedasticity. Random and fixed factors in the fitness model were identical to the resistance model: the random variable was the populations, and the three fixed effects were the treatments, replicates runs and blocks. Similarly, LSD multiple comparison test (using predictmeans package in R 3.5.0 software) were run to test the differences between treatments when significant effects of selective histories were detected by the models.

4.3 Results

4.3.1 Populations and evolutionary dynamics

A glyphosate selection experiment was conducted to examine the adaptation of *C. reinhardtii* to different rates of environmental change. For each of two final selective doses (1 and 2 MIC), three contrasting rates of change were applied: quick (i.e. sudden change selecting with the highest selective dose), intermediate (gradual doses increase in three steps up to the highest selective dose) and slow rates of change (gradual doses increase in six steps up to the highest selective dose). Cell concentrations over time for each of the seven treatments is presented in Figure 16.

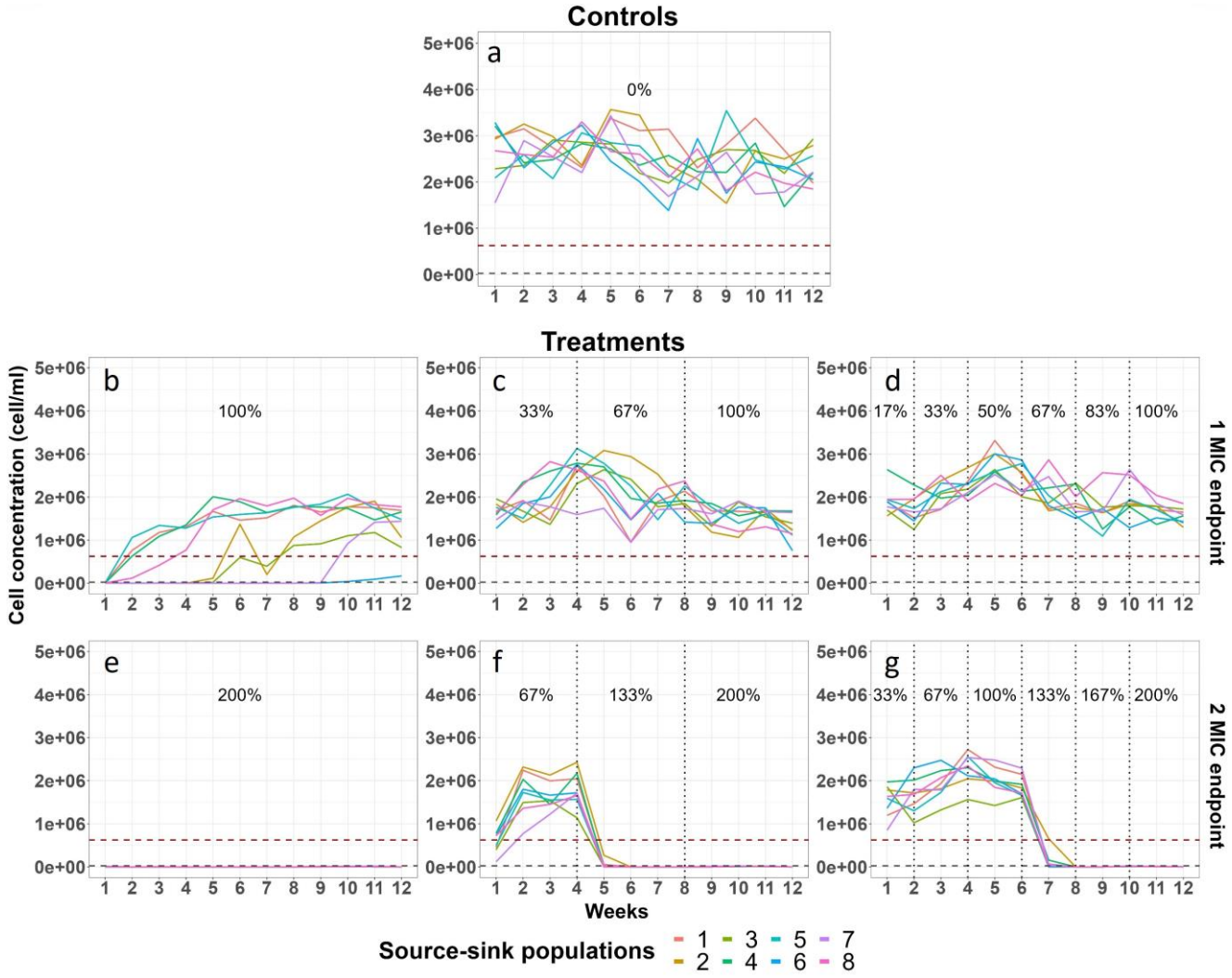

Figure 16 Evolutionary dynamics of *C. reinhardtii* source populations (a) or sink population subjected to glyphosate (b-g) at different rates of change. Doses increasing either towards 1 MIC (b-d) or 2MIC(e-g). Three contrasting rates of change were applied: quick (1 dose), intermediate (3 doses) and slow (6 doses). The cell density of the eight replicates source-sink populations was recorded over 12 weeks. Grey dash lines represent the detection limit. Red dash lines represent the threshold cell density over which populations are considered resistant. The selective doses applied are expressed in % inhibition of growth induced in a control population (see Table 13).

After the first week of selection, glyphosate doses 0mg/l, 65mg/l, 80 mg/l,115mg/l, 160mg/l and 320 mg/l resulted in the desired 0%,17%, 33%,67% and 100% inhibition of cell growth (Figure 16).

When selecting populations up to 1 MIC with intermediate and slow glyphosate dose increases (Figure 16 c and d), average cell concentrations were higher than expected when first exposed to the next doses (Figure 16 c: weeks 4 and 8, Figure 16 d: weeks of even numbers). All populations subjected to gradual rates of change maintained high populations cell densities and never required immigration from their source populations to avoid a bottleneck effect at weekly transfer. Conversely, in sink populations that were not pre-exposed to lower glyphosate doses and subjected to quick rates of change (Figure 16 b), cell densities took longer to reach high population cell densities. Each sink population initially received immigration from their respective source population until population densities were high enough. Reaching a cell concentration over the 625,000 cells $.ml^{-1}$ threshold occurred at different times for each sink population except for replicate six whose growth remained strongly inhibited throughout the entire time course.

Whereas all populations exposed to 1 MIC at the end of the 12-week experiment had survived, growth was no longer detectable in populations subjected to doses exceeding 133% inhibition regardless of the rate of environmental change (Figure 16 e, f, and g). These populations were therefore excluded from subsequent fitness and glyphosate resistance assays.

4.3.2 Evolved resistance to glyphosate

The 32 populations that survived the selection experiment were tested to confirm that they had evolved resistance to glyphosate, and to determine if selection histories affected their level of resistance. All populations were first kept for 1 week in their ancestral environment (i.e. BM) before being assayed at 1 MIC.

The cell density of each assay was measured during the exponential growth phase (after seven days), allowing an estimate of the effect of glyphosate selection regimes on growth rates. As expected, growth of controls exposed to 1 MIC was completely inhibited and led to a zero-inflated data set. In order to cope with this constraint and investigate glyphosate resistance in the glyphosate selected regimes, the values of the controls were excluded from the LMM analysis (Table 14).

Table 14: ANOVA tables from the LMM analysis run to assess the effect of selection histories on glyphosate resistance in the populations undergoing the three selection regimes (quick, intermediate and slow) during the exponential growth phase (7 days). The experiment was divided into two separate runs to fit in the incubator which was divided in blocks to account for a gradient affecting growth. Both the effects of runs and block were accounted for in the analysis.

General Linear Models ANOVA tables

Levels of resistance

Traits	Variables	SS	MeanSS	Numdf	Dendf	Fvalue	p values	
Exponential growth	Selection regime	0.46	0.23	2	21	0.46	0.64	
	Run	2.84	2.84	1	64	5.66	0.02	*
	Block	0.07	0.07	1	71	0.14	0.71	

While no growth was detected in any control population, glyphosate selected populations were confirmed to have evolved resistance, and reached an average cell concentration of 507,450 cells ml^{-1} When comparing the effect of the selection regimes, there were no significant differences (Table 14) between the average cell concentration of populations previously undergoing one of the three rates of selection (Table 15) . Previous selection with glyphosate conferred similar levels of resistance to glyphosate at 1 MIC regardless of the rate at which increase in selective dose was applied.

Table 15 : Average cell concentration at 1 MIC of glyphosate resistant populations with different selective histories after seven days in BM. Controls were never subjected to glyphosate selection while other populations underwent sudden (quick), or gradual (intermediate and slow) glyphosate-dose increase up to 1 MIC as described in the selection experiment.

Average cell concentrations at 1 MIC

Selective histories	Average cell densities (cell/ml)	Standard deviation	n
Control	-8607	11326	43
Quick	520916	421810	43
Intermediate	468732	342433	47
Slow	409565	354868	40

4.3.3 Cost of glyphosate resistance

To determine the effect of selection histories on costs associated with resistance, the fitness of all 32 populations that survived the selection experiment was estimated in their ancestral environment (i.e. BM) through the measure of growth rate. The cell density of each population was measured during the exponential growth phase (after seven days) and at carrying capacity (after 14 days).

Table 16: ANOVA tables from the LMM analysis run to assess the effect of selection histories (control, quick, intermediate and slow) on fitness in ancestral media (BM) during population's exponential growth rate phase (7 days) and at carrying capacity (14 days).

Linear Mixed Models ANOVA tables

Fitness in ancestral media

Traits	Variables	SS	MeanSS	Numdf	Dendf	Fvalue	p values	
Exponential growth	Selection regime	0.12	0.04	3	27	5.44	0.005	**
	Run	0.05	0.05	1	86	7.5	0.01	**
	Block	0.3	0.3	1	102	40.95	<0.001	***
Carrying capacity	Selection regime	0.01	0	3	27	0.99	0.41	
	Run	0.05	0.05	1	87	11.05	0.001	**
	Block	0	0	1	106	0.42	0.52	

Population cell density in the ancestral environment was significantly different between the selection histories during exponential phase ($F_{3,27} = 5.44$, $P<0.01$), but no longer at stationary phase (Table 16). During the exponential growth phase (Figure 17.a), populations that experienced a gradual selection (i.e. slow and intermediate dose increases) had a lower cell density than populations that experienced a sudden dose increase. Populations subjected to slow rates of change also grew slower than the controls.

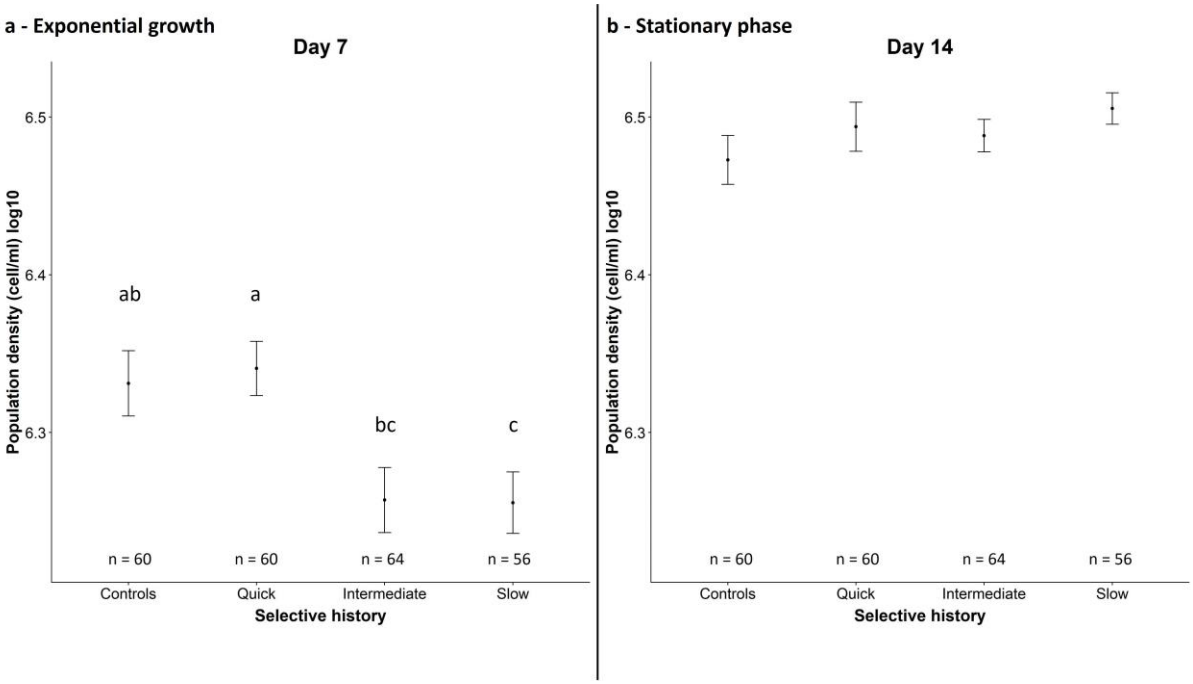

Figure 17 : Fitness in ancestral media of populations with different selective histories.Cell densities at seven and fourteen days were taken as a measure of growth rate used as a proxy for fitness. Controls were never subjected to glyphosate selection while other populations underwent sudden (quick), or gradual (intermediate and slow) glyphosate-dose increase up to 1 MIC as described in the selection experiment. Averaged cell concentrations of populations during exponential growth (a) and stationary phase (b) are represented with their standard errors. Letters represent significant differences between selective histories (LSD multiple comparison tests) when significant differences between treatments were established with the LMM analysis (see Table 14)

4.4 Discussion:

I investigated how the rate of directional environmental change affects the evolutionary dynamics and outcome of glyphosate resistance selection using experimental evolution in *C. reinhardtii*. More specifically, I tested how the rate of glyphosate dose increase affects the evolution of glyphosate resistance over the course of a 12-week selection experiment, and if there was an effect of the different selection histories on adaptation to glyphosate and potential associated fitness costs. To achieve this, I tested the following predictions:

(i) Rapid increases in glyphosate doses will reduce population size, increase variability and inconsistency of evolutionary dynamics and delay evolution of resistance.

(ii) Gradual glyphosate dose increase will allow for evolution of resistance to higher doses

(iii) Selective histories will affect fitness in ancestral environment: Adaptation to Gradual glyphosate dose increase will yield higher fitness cost.

Results support the first hypothesis and are consistent with findings in former studies in *C. reinhardtii* (Collins and De Meaux, 2009), in bacteria (Perron et al., 2006, 2008) and in simulations (Collins et al., 2007). At the onset of selection, a sudden increase in glyphosate dose through direct exposure to 1 MIC drastically decreased population size. Immigration from their source populations (controls) was required to avoid extinction and allow adaptation to occur. Additionally, there was a lot of variability in the emergence of glyphosate resistance between replicates: time to obtain consistent growth in these populations varied considerably, with one population requiring as little two weeks, and another never evolving resistance over the experimental time-course. Conversely, population size remained more stable throughout selection with gradual increase towards one 1 MIC. Populations did not risk extinction despite exposure to glyphosate and therefore did not require immigration. Being previously exposed to lower doses, population densities under gradual selection were not strongly impacted when first exposed to 1MIC as would a naïve population.

Populations that were initially sensitive evolved glyphosate resistance under the three selection regimes (quick, intermediate, or slow). The assay conducted after the selection experiment confirmed that glyphosate resistance to 1 MIC is heritable for populations subjected to any one of the three selective histories. This implies that evolution of resistance to glyphosate occurred through adaptation by acquisition of mutation(s) and/or heritable epigenetic modifications, rather than through acclimation or other means. Heritable epigenetic modifications have been identified in *Chlamydomonas* (Kronholm et al., 2017), but model predictions indicate that they may occur at earlier stages of adaptation and be later replaced by mutations (Kronholm and Collins, 2016). Glyphosate resistance acquired through adaptation in *C. reinhardtii* is consistent with previous work in this species (Lagator, 2012c; Lagator et al., 2014a, 2014b; Melero-Jiménez et al., 2021).

Although the current study did not strictly address high-dose versus low-dose selection here, some of the evolutionary principles around selection at different dose-rates are likely applicable. Higher doses impose more intense selection pressure on mutations that confer resistance (Costelloe et al., 2010; Day & Read, 2016; Kouyos et al., 2014), and as such it is widely suggested that lower doses are preferrable to slow the spread resistance (Blanquart, 2019; Kouyos et al., 2014). Some studies in higher plants however have shown that even very low doses of herbicide can rapidly select for resistance, thought to be via the rapid assembly of polygenic resistance traits (Never and Powels, 2005). The results presented here may be interpreted according to evolutionary rescue theory: we can speculate that populations avoided extinction under sudden changes in glyphosate dose through mutations of sufficiently large effect (Perron et al., 2008) either arising *de novo* or imported through immigration. In the context of this study, the occurrence or evolutionary rescue of *C. reinhardtii* populations exposed to rapid glyphosate dose increases is delayed and less predictable than under gradual glyphosate dose increases. Accordingly, we can speculate that under rapid changes adaptation may be mutation limited. Conversely, under gradual change in glyphosate doses we speculate that multiple different mutations could convey sufficient resistance to this weaker selection pressure and that a greater number of possible evolutionary trajectories exist. Over time, iterative stacking of multiple mutations would increase glyphosate resistance leading to individuals carrying a greater number of more varied mutations.

While hypothesis one was broadly supported, results do not support the second prediction. Gradual rates of changes did not facilitate an expansion of *C. reinhardtii*'s niche to doses above 1MIC. A moderate or slower increase in glyphosate dose did not allow populations to evolve resistance to doses higher than 215

mg l^{-1} (1.3 MIC): growth of all populations remained zero regardless of immigration from the source populations. These results are similar to those observed by Hansson et al. (2024) using the same CC-1690 *C. reinhardtii* strain, where populations never grew at doses above 1 MIC (Hansson et al., 2024). From these results, one can speculate that adaptation to glyphosate doses above 1 MIC may require even slower rates of change, larger population size or may remain an unattainable adaptive peak. Another study used a ratchet protocol to expose a *C. reinhardtii* lake isolate to increasing doses of glyphosate. The dose would be increased only when the growth rate of the selected population reached the growth rate of control populations. In this case the rate of change is slower than the ones applied in our study as it gives the opportunity for larger populations to arise before increasing selective pressure. Authors observed that the initial MIC (90ppm) could be shifted to 1.8 MIC (160 ppm) (Melero-Jiménez et al., 2021). This seems to suggest that adaptation to higher selective pressure may be obtained if populations are large enough, and if the rate of change is slow enough. However, since this experiment was conducted with a different strain, there may also be an impact of the genetic background.

Lastly, it was hypothesised that higher fitness costs will be associated with adaptation to gradually increasing glyphosate doses. The results of the current experiment support this prediction: fitness in the ancestral media of populations previously adapted to slow glyphosate dose increase was significantly lower than that of the control populations. One possible explanation for this is that under gradual change, a large number of mutations of smaller effect are likely to be fixed for each local optimum (Collins et al., 2007; Perron et al., 2008) and their additive or synergetic associated costs could lead to a higher associated fitness cost (in the absence of compensatory mutations). Similar results have been found across other studies. For example, a slow increase in glyphosate dose (applied using a ratchet protocol) led to a major fitness cost in one prior study of *C. reinhardtii* (Melero-Jiménez et al., 2021), whilst *C. reinhardtii* under selection with glyphosate directly applied at 1 MIC evolved resistance but with no associated fitness cost (Hansson et al., 2024; Lagator et al., 2012; Vogwill et al., 2012). In one case, the intermediate glyphosate dose of 0.5 MIC was also tested and yielded the same results (Hansson et al., 2024). Interestingly, in studies investigating cross-resistance, high levels of glyphosate resistance with no associated fitness cost in *C. reinhardtii* were also associated with absence of cross resistance, which suggests that glyphosate resistance in these cases may be endowed by a TSR mechanism (Lagator et al., 2012; Vogwill et al., 2012). This remains speculation since the mechanism of glyphosate resistance in *C. reinhardtii* is still unknown to date.

While the results of the current glyphosate study are well supported within the literature, there are observations of contradictory responses in relation to selection with other substrates. Selecting *C. reinhardtii* with slower rates of change in decreasing phosphate concentrations resulted in better adapted outcome and a lower fitness costs (Collins and De Meaux, 2009). While the authors also suggest that adaptation to slower rates of change are likely explained by the accumulation of relatively small effect mutations, they speculate that these resultant genotypes might have lower levels of pleiotropy and historical constrains, leading to lower costs. Two studies from Gorter et.al (2016, and 2017) used known yeast phenotypes under selection with three different heavy metals. Cadmium, nickel and zinc have different biological properties and roles in yeast, allowing to study how the nature of the selection pressure affects evolutionary dynamics and outcomes. Their results suggest that evolutionary dynamics and outcomes may depend on the selective agent used to apply the sudden or gradual rate of change, a finding echoed by Lagator et.al (2012) when investigating another type of temporal variation: herbicide cycling at different rates. In another study, *C. reinhardtii* populations were exposed to a new herbicide mode of action

once resistance to the previous one was attained, to address the impact of selection history on adaptation to multiple herbicides (Lagator et al., 2014a). Their results demonstrated a uniform decrease in fitness cost as resistance mechanisms accumulated in selective histories increasing environmental heterogeneity. Additionally, loss of resistance was never observed. According to the authors, the most plausible explanation is antagonistic epistasis between fitness costs associated with resistance to different herbicides. They speculate that sequential evolution of resistance to multiple herbicides may result in optimization of defence metabolic pathways (NTSR) reducing pleiotropic costs.

Contrasting results over several studies in *C. reinhardtii* suggest that fitness costs associated with varying rates of adaptation is not universal. This conclusion is also starting to become a consensus with regards to evolved herbicide resistance in higher plants. A review showed that the expression of a fitness cost associated with herbicide resistance was specific to the resistance gene, allele, and genetic background and species (Vila-Aiub et al., 2009). A second review focused on TSR glyphosate resistance associated fitness costs in higher plants. Single, double, and triple mutations in *EPSPS* are reported to have different associated costs (Vila-Aiub et al., 2019). Their general conclusion was that in the case of *EPSPS* mutations, identification of their effect on the trade-off between glyphosate and PEP binding, as well as studying enzyme kinetic is useful to predict glyphosate resistance associated fitness cost. In the case of amplification or overexpression of *EPSPS*, the observed results in the literature challenge the resource allocation theory (the notion that upregulation of alleles for resistance diverts resources away from plant growth or development), since there have been reports of both presence and absence of fitness cost as well as presence of a fitness benefit of glyphosate resistance. Additionally, even when a fitness cost is predicted, it may not be observed if it is compensated by genome modifications or if it is an ecological-mediated cost (i.e. when the fitness cost is expressed in presence of biotic or abiotic stress). The same principle could apply to NTSR mechanisms as demonstrated by a study detecting a reproductive cost of NTSR mechanism in *Alopecurus myosuroides* only under nutrient deprivation (Comont et al., 2019b). The authors propose that evolved resistance cause variation in developmental traits which in interaction with other stressors can result in indirect fitness cost.

Regarding higher fitness costs of glyphosate adaptation being associated with gradual rates of change, one potential explanation may be the initial difference in selective pressure resulting in smaller populations size under sudden rates of change. Smaller population sizes reduce the probability of a resistance mutation arising in the sink populations and increases the likelihood that few fixed mutations arise when, by chance, a mutation of sufficiently large effect with a low associated cost is sampled from the source populations kept in the ancestral environment (Perron et al., 2008). When glyphosate dose increases gradually, population sizes remain large enough for mutations to arise in the selected population without the need for immigration (Perron et al., 2008). As selective pressure increases gradually, fitness is repeatedly increased by small amounts by mutations of smaller effects that are likely to be fixed for each of these local optimums (Collins et al., 2007; Perron et al., 2008). As these mutations arise within populations under continual glyphosate exposure, any potential variation in their fitness in the ancestral environment is likely outweighed by their selective advantage in the presence of glyphosate. As a result, the higher cost associated with gradual changes could be explained by fixation of a larger number of mutations and their additive associated cost in absence of compensatory mutations. It is currently outside the scope of this study to test this hypothesis, but it could be investigated in further work.

In conclusion, within this study it was identified that rapid increases in glyphosate doses increased both variability and inconsistency in evolutionary dynamics and delayed evolution resistance. Additional findings

were that adaptation to gradual glyphosate dose increases incurred fitness cost and did not allow *C. reinhardtii* populations' adaptation to doses above 1 MIC. One important consideration for further interpretation of these results is to understand more explicitly the genetic basis for resistance which has evolved under these different selective environments. The understanding gained with the investigation of phenotypes observed during the selection experiment could be further completed by understanding the underlying genetic basis of adaptation and to attribute identified pleiotropic effects to specific genes and mutations, as well as for understanding their biochemical and physiological origins and underlying causes (Vila-Aaiub et al., 2009). Consequently, in the next chapter I propose to investigate the genetic basis of glyphosate resistance in *C. reinhardtii* to better understand the molecular mechanisms and their effects on the fitness cost.

5 Investigating the impact of glyphosate selective history on genomic changes in *C. reinhardtii*

5.1 Introduction

In nature, organisms typically face constant environmental changes. For populations to persist in the long term under such conditions, they must either migrate to more favourable habitats or adapt phenotypically or genetically *in situ* (Burger and Lynch, 1995). One particularly relevant example of environmental change is directional change, exemplified by global climate change and human-induced pollution. The rate at which such directional environmental changes occur can significantly impact population persistence (Gonzalez and Bell, 2013; Lindsey et al., 2013) and influence evolutionary dynamics and outcomes (Collins and De Meaux, 2009; Toprak et al., 2012).

Despite this, most experimental evolution studies focus on adaptation to a sudden, large environmental shift, after which the environment remains stable (Barrick and Lenski, 2013; Collins, 2011). There are, however, several theoretical predictions about how the rate of environmental change influences evolutionary processes and outcomes. Many of these predictions fall within the framework of the "moving optimum model", which considers adaptation to directional environmental change as the evolution of a quantitative trait under stabilizing selection with a shifting optimum (Kopp and Hermisson, 2007, 2009b, 2009).

According to this model, gradual environmental changes are predicted to drive adaptation through the repeated fixation of mutations with intermediate phenotypic effects, whereas sudden large changes are expected to result in the fixation of fewer mutations with relatively larger phenotypic effects. Since mutations with smaller phenotypic effects are more common, gradual changes may lead to more diverse evolutionary pathways. Furthermore, weaker genetic interactions among mutations with smaller effects (Schenk et al., 2013; Schoustra et al., 2016) suggest that evolution may be less constrained under gradual change. This, in turn, may enable populations adapting to gradual environmental shifts to achieve higher fitness levels (Collins and De Meaux, 2009).

I previously conducted experimental selection for 12 weeks (approximately 120 generations in the control populations) in *Chlamydomonas reinhardtii* exposed to varying rates of directional environmental change (Chapter 4). The treatments in the Variable Rates (VR) selection experiment involved two final selective doses (1 MIC or 2 MIC) with three rates of glyphosate dose increase: quick (1 dose), intermediate (3 doses) and slow (6 doses)). No survival was observed as soon as selective dose was above 1 MIC (final selective dose 2MIC). For surviving populations under selection up to 1 MIC, the contrasting selection regimes resulted in different evolutionary dynamics and outcomes: rapid glyphosate dose increase led to variability in evolutionary dynamics and was not associated with a fitness cost while gradual increase supported more stable population growth and evolution of resistance but resulted in significant fitness cost under the slow rates of change. Populations evolved heritable glyphosate resistance indicating that adaptation involved mutations and/or heritable epigenetic modifications.

Here, I investigate how each of the selective histories from the VR experiment affected the genetic basis of glyphosate resistance, more specifically investigating mutations underpinning glyphosate resistance under three different selective histories.

The genetic basis of glyphosate adaptation is already well characterised in higher plants (Shaner et al., 2012). Glyphosate inhibits the gene encoding for the EPSPS enzyme (Steinrucken and Amrhein, 1980). To date, target-site mutations, gene copy number variation, *EPSPS* increased transcription and EPSPS increased enzymatic activity are reported as TSR mechanisms of glyphosate resistance in higher plants (Baerson et al., 2002; Molin et al., 2017; Zhang et al., 2015). Known NTSR glyphosate resistant mechanisms reported in the literature include reduced uptake and translocation, enhanced metabolic degradation and glyphosate detoxification (Deng et al., 2022; Michitte et al., 2007; Pan et al., 2019; Vila-Aiub et al., 2011).

Using whole genome sequencing and variant calling, I analysed the genome of single clones isolated at the final time point of the VR experiment to compare evolved *C. reinhardtii* under quick, intermediate and slow selective histories. Focussing on the effect of selection histories on the number and type of mutations selected as well as the repeatability of adaptation.

5.2 Methods

5.2.1 Samples, library preparation and sequencing

The samples used here were from the Variable Rate (VR) selection experiment presented in Chapter 4. An isogenic culture was used to start an experiment in which *C. reinhardtii* was subjected to three regimes of glyphosate dose increase: quick, intermediate and slow rate of change. Populations surviving after 12 weeks of selection (see Chapter4, Figure 3-1) were stored on BM agar slopes (see Chapter 2 section 1). Unexpectedly, some of the stored populations experienced mortality only 4 months after storage. Most of the populations from replicates 1, 4, 5 and 8 (with exception of the replicate 1 with a slow selective history) and the isogenic population used to inoculate the VR experiment (T0) survived and were transferred to liquid medium and cultured. Isogenic axenic cultures (see Chapter 2, section 4.1 for detailed protocol) were prepared for cell harvest and DNA extraction (see Chapter 2 section 4.2)

The 16 cultures were also checked for contamination on nutrient agar plates which confirmed them to be axenic. Cells from all 16 cultures were harvested for DNA extraction and DNA was extracted using an in-house DNA extraction protocol for Chlamydomonas (for details see Chapter 2 section 4.3 and Appendix 2). DNA integrity was assessed by agarose gel electrophoresis and purity determined using a nanodrop spectrophotometer. Quantity was determined by fluorometric quantification using the Qubit dsDNA BR Assay Kit. PCR-free library preparation and sequencing was performed on all 16 samples by the Center for Genomic Research (Liverpool) using an Illumina NovaSeq sequencer which produced 150bp paired-end reads and an average coverage of 200X per sample. The VR experiment sample names and selective histories are as detailed below (GS = glyphosate-sensitive , GR = glyphosate-resistant, replicates D to G):

Selective history	Sample names:
Controls	GSD
	GSE
	GSF
	GSG
Quick	GR-quickD
	GR-quickE
	GR-quickF
	GR-quickG
Intermediate	GR-intermD
	GR-intermE
	GR-intermF
	GR-intermG
Slow	GR-slowE
	GR-slowF
	GR-slowG

5.2.2 Variant calling and genotyping

General steps, tools and their parameters for the variant calling and genotyping pipeline are described in detail in Chapter 2. These were equivalent to the finalised pipeline used for analysis of the SR experiment described in Chapter 3.

5.2.2.1 Read processing

Prior to delivery, raw reads were processed by the Center for Genomic Research as follows : Illumina TruSeq2-PE adapter sequences were removed before reads ends were further trimmed using a window size of 5 and minimum quality score of 20. Reads shorter than 50 bp after trimming were removed. Post trimming read assessment (FastQC v0.11.9) revealed high overall paired-end read quality, with Phred scores > 34 across the read length. GC content matched the expected genome composition, and duplication levels were low. Trimmed sequencing data was deemed suitable for subsequent variant calling analysis.

5.2.2.2 Alignment and variant calling

The reference genome used was the a highly contiguous nanopore nuclear genome assembly of the CC-1690 laboratory strain published in 2020 (O'Donnell et al., 2020) further referred to as O'Donnell-CC1690_v1.0 (for more details, see Chapter 2 section 5.1.1) in this chapter.

Table 17: Alignment metrics averaged across 16 samples from the Variable Rate (VR) selection experiment. Values were obtained from Qualimap(v2.2.1).

Alignment QC(Qualimap):	average
Mean read length	143.41
Reads mapped (%)	99.87
Properly paired reads (%)	99.805
Mean coverage	137.52
Standard deviation coverage	212.24
Mean mapping quality	51.90

Variant calling results were further filtered (see Chapter 2 for detailed information on filters) to retain high quality variants that may be linked to glyphosate resistance (Table 18). Filtered variant lists for pairs in the slow selective history were identical, posing questions regarding the reality of these results. On further investigation, there was evidence that an error occurred in the early step of the variant calling pipeline affecting the BAMs from the slow selective history. To allow for selective histories comparison, all samples must be processed by Freebayes simultaneously, ensuiring that local realignments around indels are identical in every sample (Ness et al., 2012). Due to time constraints, troubleshooting and re-running the analysis before submission of this thesis was not possible. Consequently, I choose to exclude the slow selective history results for the time being. For an annotation, the previously transferred (Chapter 3) Phytozome v5.6 annotation (using Liftoff v1.6.3) was used. Variant annotation was then performed with the newly annotated O'Donnell-CC1690_v1.0 (SnpEff v4.3+T.galaxy2).

Table 18: Number of variants called in each replicate GR/GS pair of the VR experiment at different false positive filtering steps for each of the three selective histories in Chapter 4 (quick, intermediate and slow rates of glyphosate dose increase). The samples presented here come from the VR selection experiment where one isogenic sensitive population was used to inoculate all environments. Ancestral environment (generating GS samples) and glyphosate selective environments with 1 MIC as the highest dose (generating GR samples). Three rates of glyphosate dose change were applied over 12 weeks. Quick change was imposed by selecting immediately with the endpoint dose, intermediate (three doses) and slow (six doses) change by gradual increase of doses towards the endpoint. Populations were transferred weekly into fresh media. DNA samples presented in this chapter come from GS and GR samples that underwent the full 12 weeks of selection.

Selective histories	Quick	Intermediate	Slow
1-GS and GR replicate pair are different			
Rep D	1178	1230	NA
Rep E	1203	1310	926
Rep F	1269	1258	953
Rep G	1128	1212	863
2-GR has non-reference call			
Rep D	922	931	NA
Rep E	869	958	884
Rep F	938	926	957
Rep G	847	909	885
3-GR call is not in any GS sample			
Rep D	246	257	NA
Rep E	259	326	351
Rep F	317	311	392
Rep G	269	305	386
4-Alternate allele frequency in GR is >0.85			
Rep D	14	31	NA
Rep E	20	36	25
Rep F	32	35	44
Rep G	29	33	52

5.3 Results

5.3.1 Analysis of the EPSPS-encoding gene

As a known host of target-site resistance mutations in other organisms, the EPSPS-encoding gene (Schönbrunn et al., 2001) was investigated as previously described in Chapter 3. Firstly, there was no *EPSPS* mutation potentially linked to the glyphosate resistant phenotype: although three loci with variants were called by Freebayes in *EPSPS*, they did not pass filtering criteria as genotype did not differ between GS and GR samples in any of the selective histories. Secondly, there was no evidence for *EPSPS* copy number variation in any selective history as each aligned read counts ratios for replicate GS/GR pairs were very close to 1 (Table 19).

Table 19: *EPSPS* copy number detection in the quick,intermediate and slow selective histories. Normalised *EPSPS*-aligned read counts are estimated for all samples. GR/GS ratios are calculated by dividing the value for a glyphosate resistant sample by the value for the glyphosate sensitive sample of the same replicate (D, E, F or G).

GR/GS ratios	D	E	F	G
Quick	0.99	1.01	1.05	1.01
Intermediate	0.97	0.96	1.02	1.00
Slow	NA	0.98	1.02	0.96

5.3.2 Genome-wide analysis of variants and annotations

The filtered list of putative variants supporting glyphosate resistance comprised 95, 135, and 121 variants for the quick, intermediate and slow selective histories, respectively (see Table 18 and Appendix 5). Across the entire experiment, the average variant calling quality and genotype qualities (Table 20) were considered high and were largely consistent across the replicates. Variant alleles frequencies were close to 1 in the GR samples on average (as expected from the > 0.85 allele frequency filter used) and very low in the GS samples, on average.

The SnpEff-generated variant annotations, impact scores and gene names for the list of variants are provided in (Appendix 6)

Table 20: Average alternative allele frequencies and average genotype qualities for glyphosate resistant and glyphosate sensitive samples in each replicate of the three selective histories (quick, intermediate and slow) of the VR experiment

Selective history	Replicate	Sample	Average AAF	Average GQ
quick	D	GR	0.944	147.44
		GS	0.000	137.94
	E	GR	0.991	147.10
		GS	0.001	140.45
	F	GR	0.992	123.62
		GS	0.010	150.27
	G	GR	0.984	149.44
		GS	0.017	143.14
intermediate	D	GR	0.994	153.05
		GS	0.002	148.76
	E	GR	0.974	139.68
		GS	0.027	142.71
	F	GR	0.975	143.80
		GS	0.023	148.58
	G	GR	0.991	144.02
		GS	0.014	148.71
slow	E	GR	0.987	146.72
		GS	0.012	139.89
	F	GR	0.996	152.55
		GS	0.002	153.55
	G	GR	0.973	149.56
		GS	0.004	141.85

5.3.3 Variant description

SnpEff (v4.3+Tgalaxy2) was used to categorise and annotate the final variant list based on the predicted effects of the genetic variants (see Chapter 2 for more detail). In the quick (Figure 18), intermediate (Figure 19) and slow (Figure 20) selective histories, single nucleotide polymorphisms (SNPs), insertions and deletions were found in all four replicates, with proportions of each relatively stable across the different replicates. Complex variants were only found in replicates E, F and G.

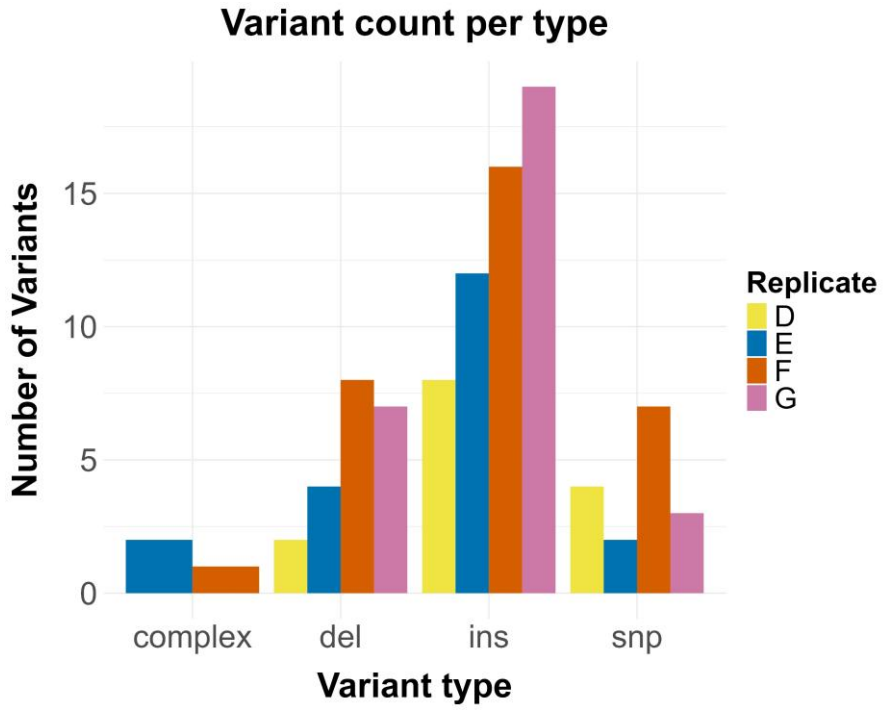

Figure 18:Variant type in each replicate of the samples with quick selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2). In snpEff, complex variants types are variants combining multiple types of change and do not fall into the simple categories of deletions (del), insertions (ins) and single nucleotide polymorphism (snp).

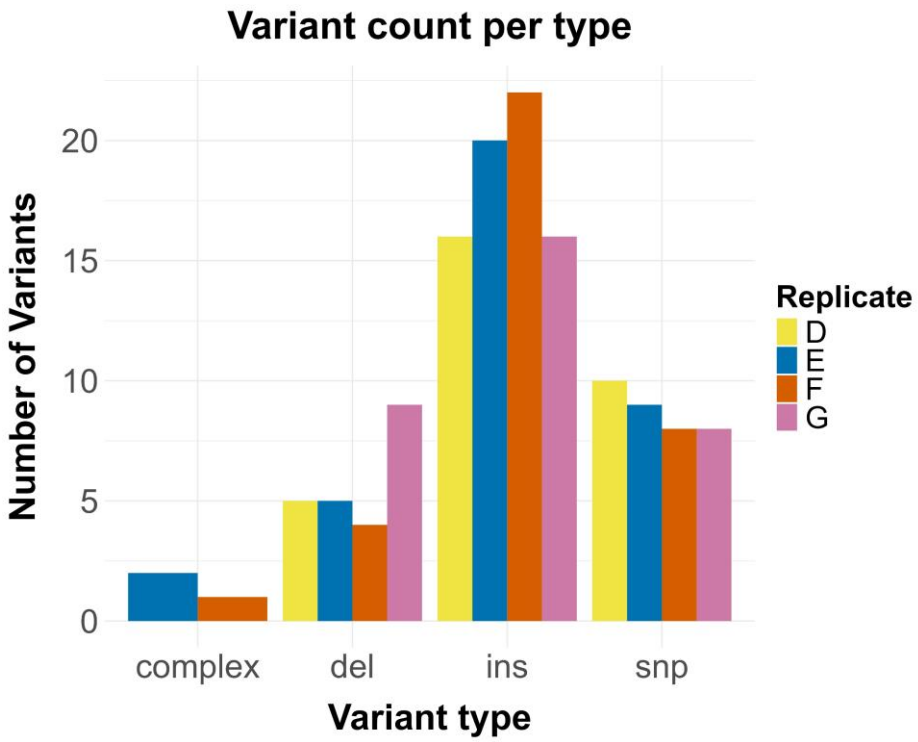

Figure 19:Variant type in each replicate of the samples with intermediate selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).In snpEff, complex variants types are variants combining multiple types of change and do not fall into the simple categories of deletions (del), insertions (ins) and single nucleotide polymorphism (snp).

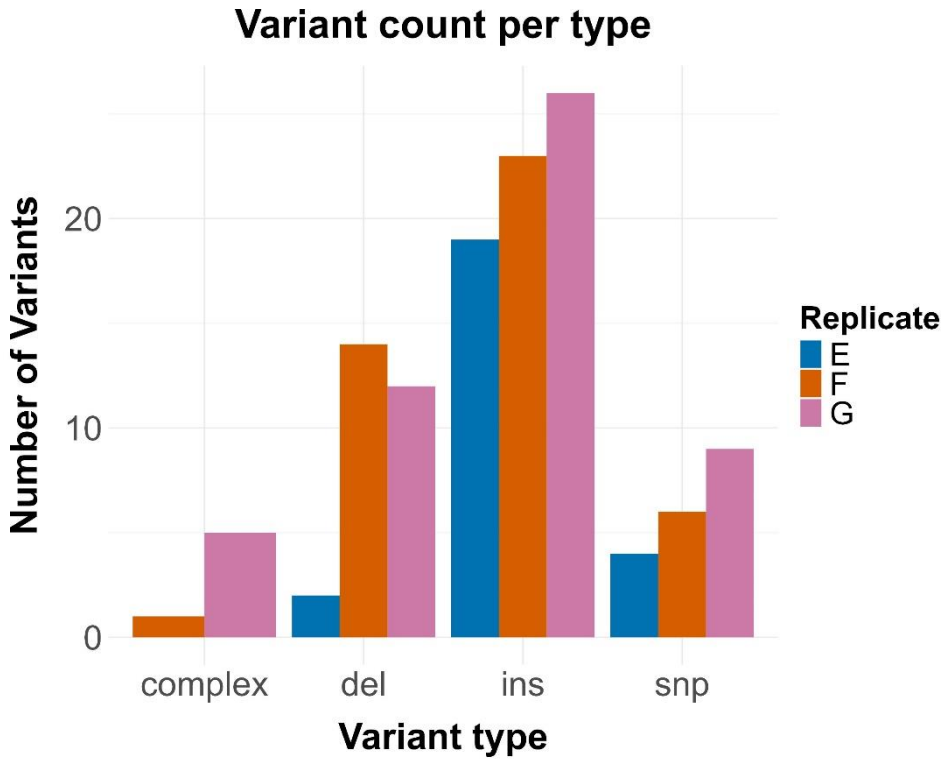

Figure 20: Variant type in each replicate of the sample with slow selective history from the VR experiment. Annotated with SnpEff (v4.3+T.galaxy2). In snpEff, complex variants types are variants combining multiple types of change and do not fall into the simple categories of deletions (del), insertions (ins) and single nucleotide polymorphism (snp).

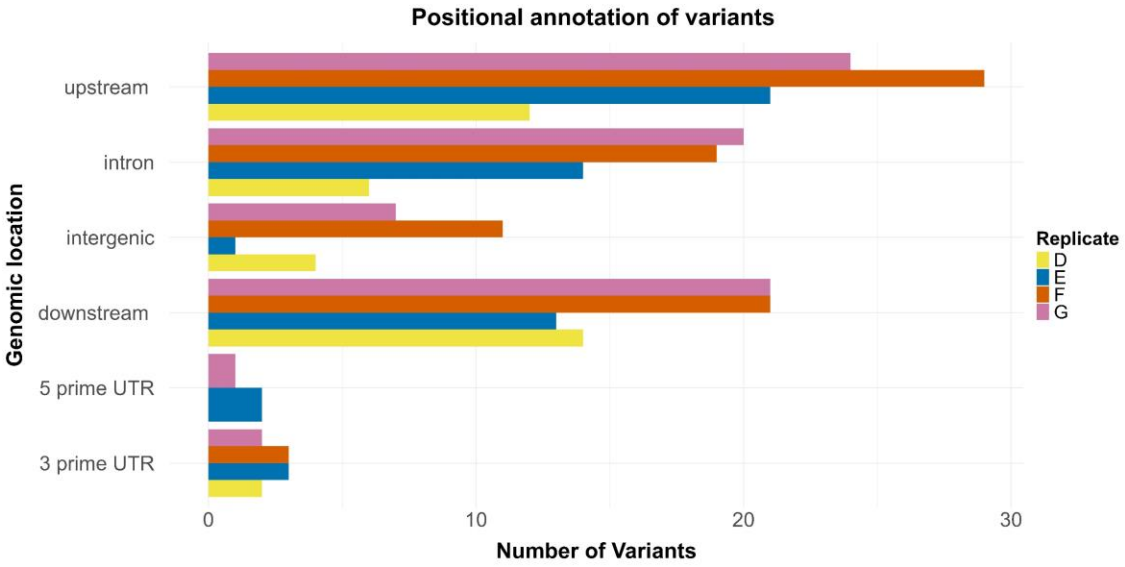

Figure 21:Variant location relative to the predicted impacted gene in each replicate of the samples with quick selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

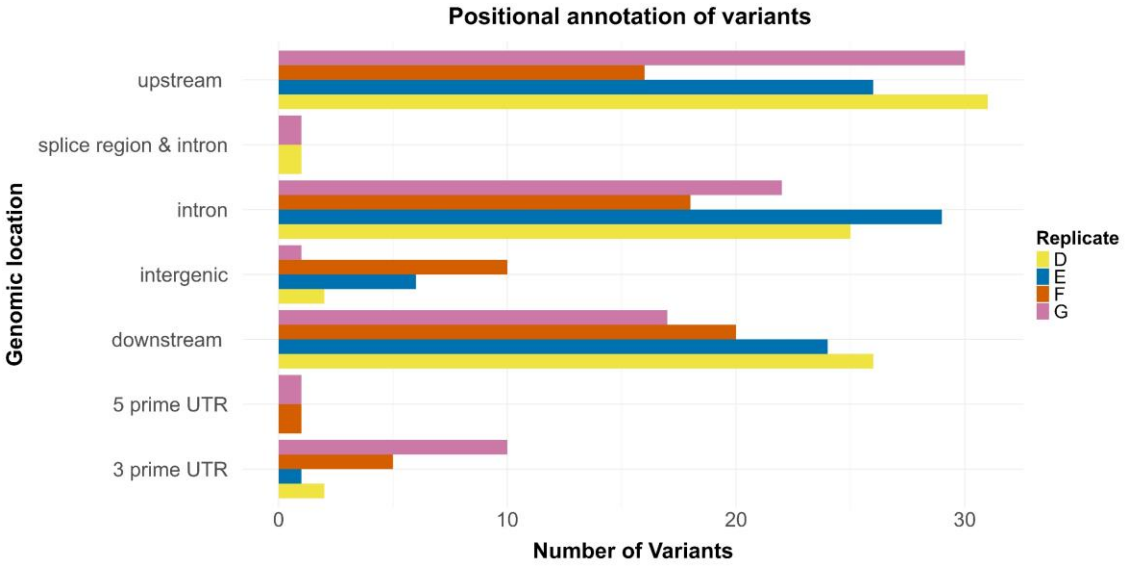

Figure 22:Variant location relative to the predicted impacted gene in each replicate of the samples with intermediate selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

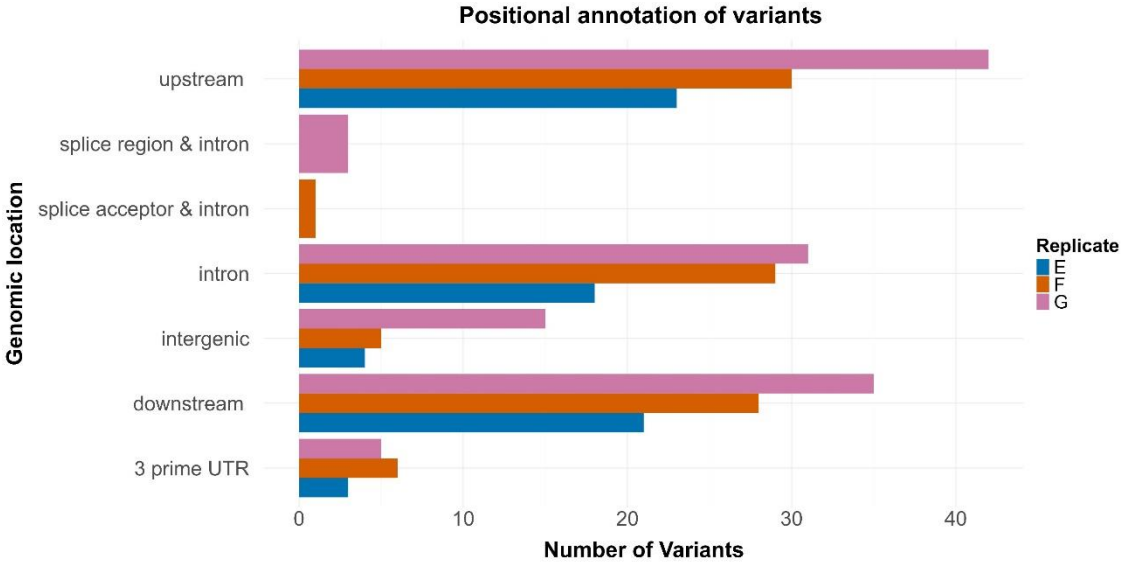

Figure 23: Variant location relative to the predicted impacted gene in each replicate of the samples with slow selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

In all three selective histories, the majority of variants detected were insertions for all replicates (Figure 18, Figure 19 and Figure 20)

SnpEff provides detailed information about the position of each variant within genomic features in the annotations.

When considering the positional annotation of variants relative to genes, the majority of variants were located in non-coding regions upstream, downstream of genes, within introns or in intergenic regions in all selective histories (Figure 21, Figure 22 and Figure 23).

SnpEff impact scores (in order of predicted impact magnitude: high, moderate, low and modifier) were used as a starting point to prioritise candidate variants. Seven variants classed as high impact were identified in the quick, intermediate and slow selective histories (Figure 24, Figure 25 and Figure 26). None of these were shared amongst replicates. The five genes that are predicted to be affected by high impact variants were assigned functional annotations (Table 21).

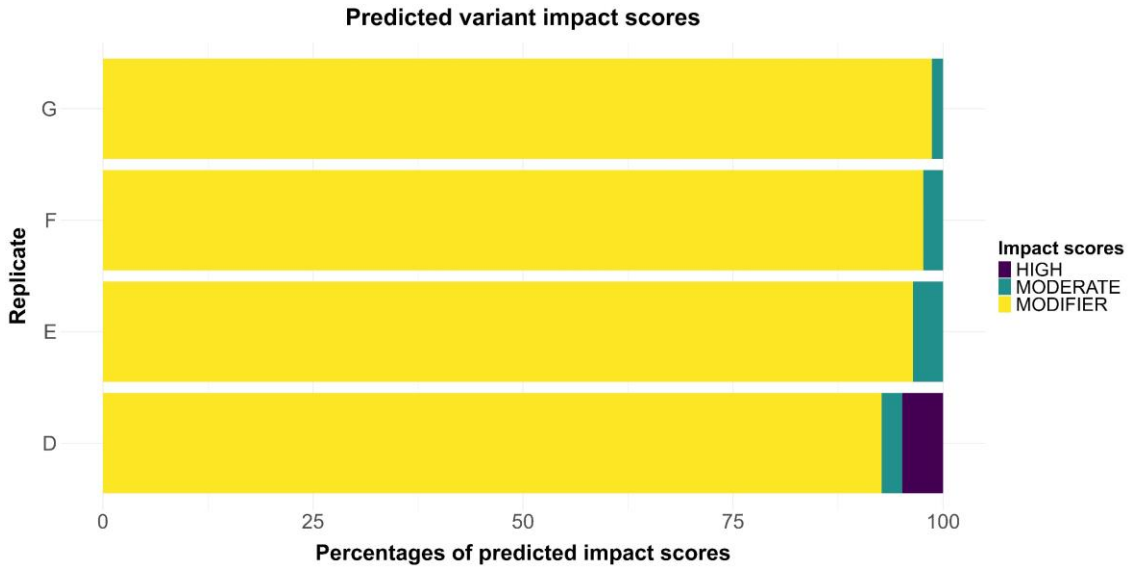

Figure 24: Predicted variant impact score in each replicate of the samples with quick selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

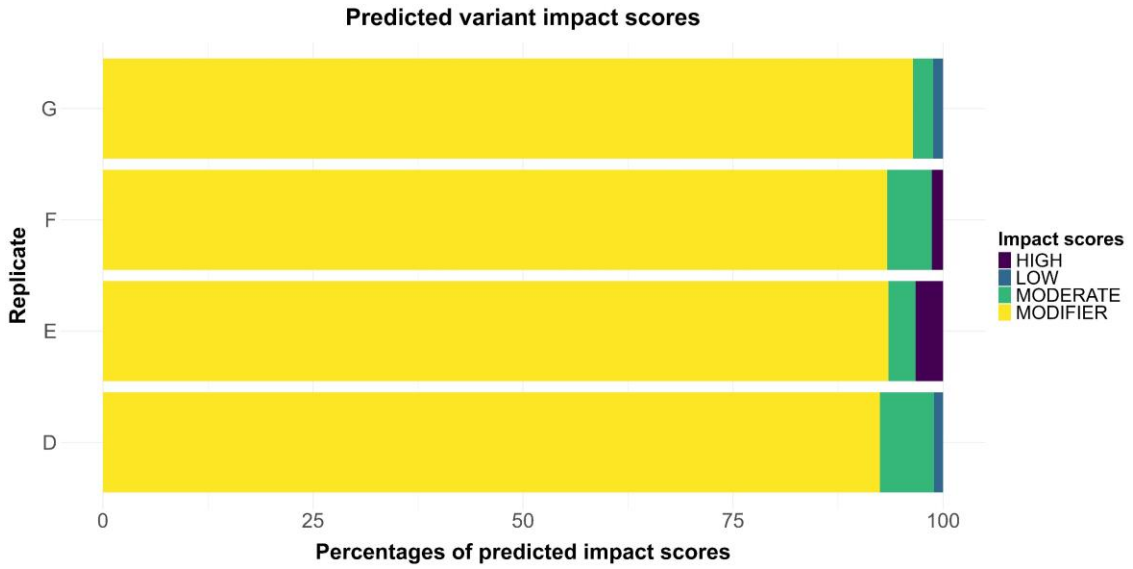

Figure 25:Predicted variant impact score in each replicate of the samples with intermediate selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

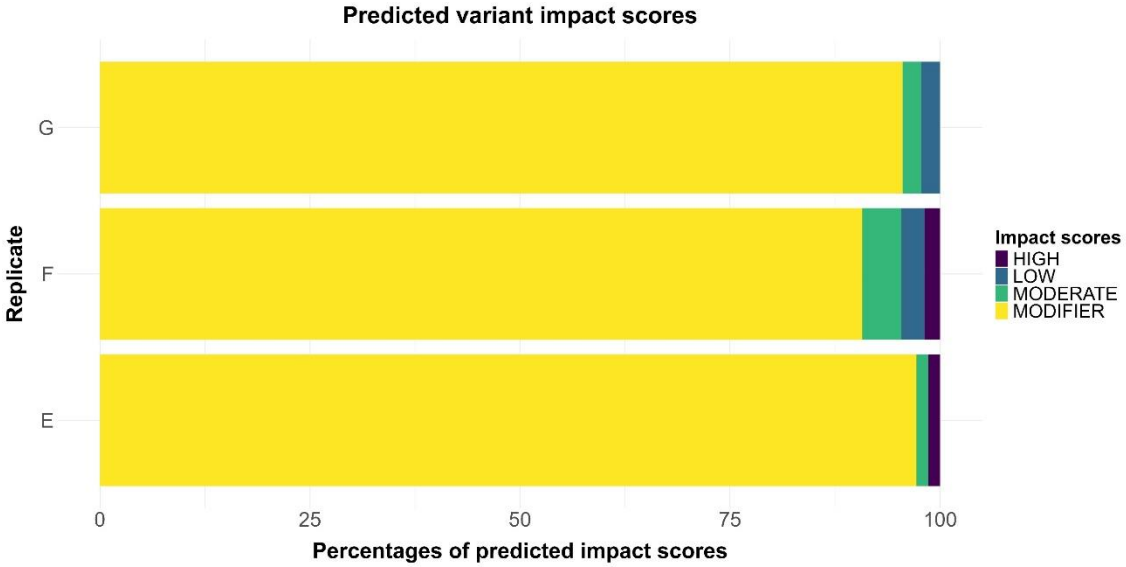

Figure 26: Predicted variant impact score in each replicate of the samples with slow selection history from the VR selection experiment. Annotated with SnpEff (v4.3+T.galaxy2).

Table 21: Variants classified by SNPeff (v4.3+T.galaxy2) as HIGH impact in the VR selection experiment for quick, intermediate and slow selective histories. Gene name associated PlantFAMS information was extracted from Phytozome (https://phytozome-next.jgi.doe.gov/).

Selective history	Chromosomes	Position	Replicate	Annotation	Gene	Associated PlantFAMS	
						Viridiplantae	Chlorophyte
quick	10	2534998	D	frameshift variant	*Cre10.g436200*	PTHR22979 ZINC FINGER PROTEIN-RELATED	PF13639 - Ring finger domain (zf-RING_2)
quick	11	813533	D	stop gained	*Cre11.g467750*	PTHR23033 – BETA1,3-GALACTOSYL-TRANSFERASE	PTHR23033 - BETA1,3-GALACTOSYLTRANSFERASE

interm	9	1209544	F	frameshift variant	*Cre09.g400950* (*NCT2*)	Permease of the major facilitator superfamily	Permease of the major facilitator superfamily
interm	10	2932776	E	splice donor & intron variant	*Cre10.g440000* (*OPR120,RAA8*)	PF08393//PF12777//PF12781 - Dynein heavy chain, N-terminal region 2 (DHC_N2) // Microtubule-binding stalk of dynein motor (MT) // ATP-binding dynein motor region D5 (AAA_9)	PTHR10015 - HEAT SHOCK TRANSCRIPTION FACTOR
interm	16	801084	E	frameshift variant	*Cre16.g647602*	SNF2 domain-containing protein / helicase domain-containing protein	Chromatin remodelling complex WSTF-ISWI, small subunit
slow	7	451071	E&F	frameshift variant	*Cre07.g315350*	TBC domain containing protein, expressed	Ypt/Rab GTPase activating protein

slow	12	702175	F	splice acceptor & intron variant	*Cre12.g492750*	3.2.1.1 - Alpha-amylase / Glycogenase

5.3.4 Genomic distribution of variants

To investigate their distribution along the genome, variants were counted in overlapping windows of 100 Kb along the genome.

No variants were reported in the scaffolds (yet to be assigned to chromosome assemblies) after filtering steps to retain only variants potentially linked to glyphosate resistance (scaffolds excluded from Figure 27, Figure 28, Figure 29, and Figure 30 to improve visualisation of variants). Variants were largely distributed across the genome (Figure 27, Figure 28,Figure 29, and Figure 30) with no obvious hotspots. Comparing replicate samples within selective histories, shared 100 Kb windows containing variants in all replicates were located on chromosomes 3 and 9 in the quick selective history (Figure 27), on chromosome 10 for the intermediate selective history (Figure 28) and on chromosome 16 for the slow selective history(Figure 29).

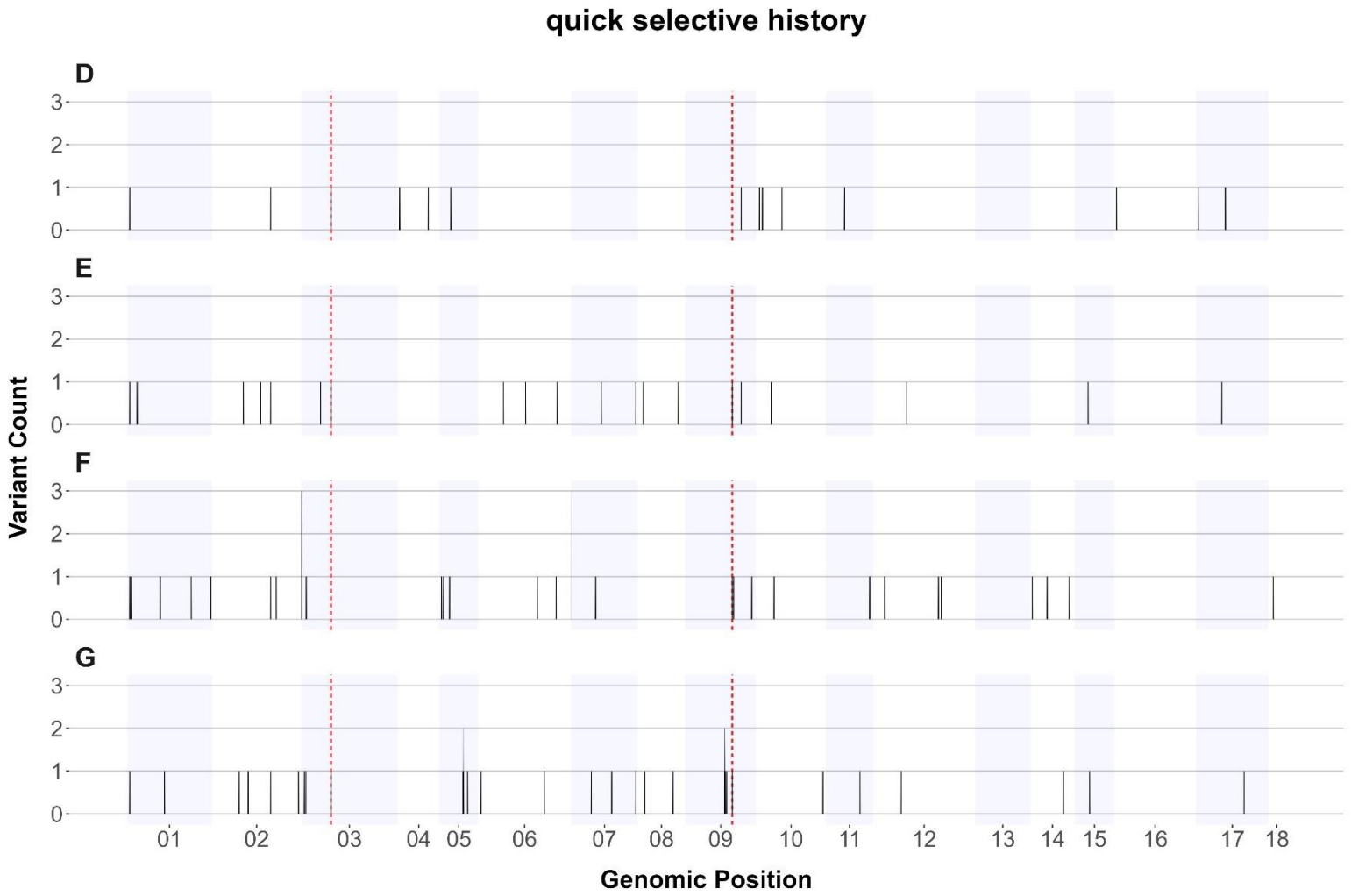

Figure 27: Genome distribution of variants in each of the replicates from the quick selective history of the VR selection experiment. Variants were counted in overlapping bins (100 Kb sliding window with a 10 Kb step) along the genome. Shaded areas delimit chromosome boundaries. Vertical black lines represent bins with variants: their thickness reflects the number of consecutive bins in a given genomic region while their height reflect the number of variants per bin (or cluster of bins). Vertical red dashed lines represent regions with bins containing variants in all three replicates.

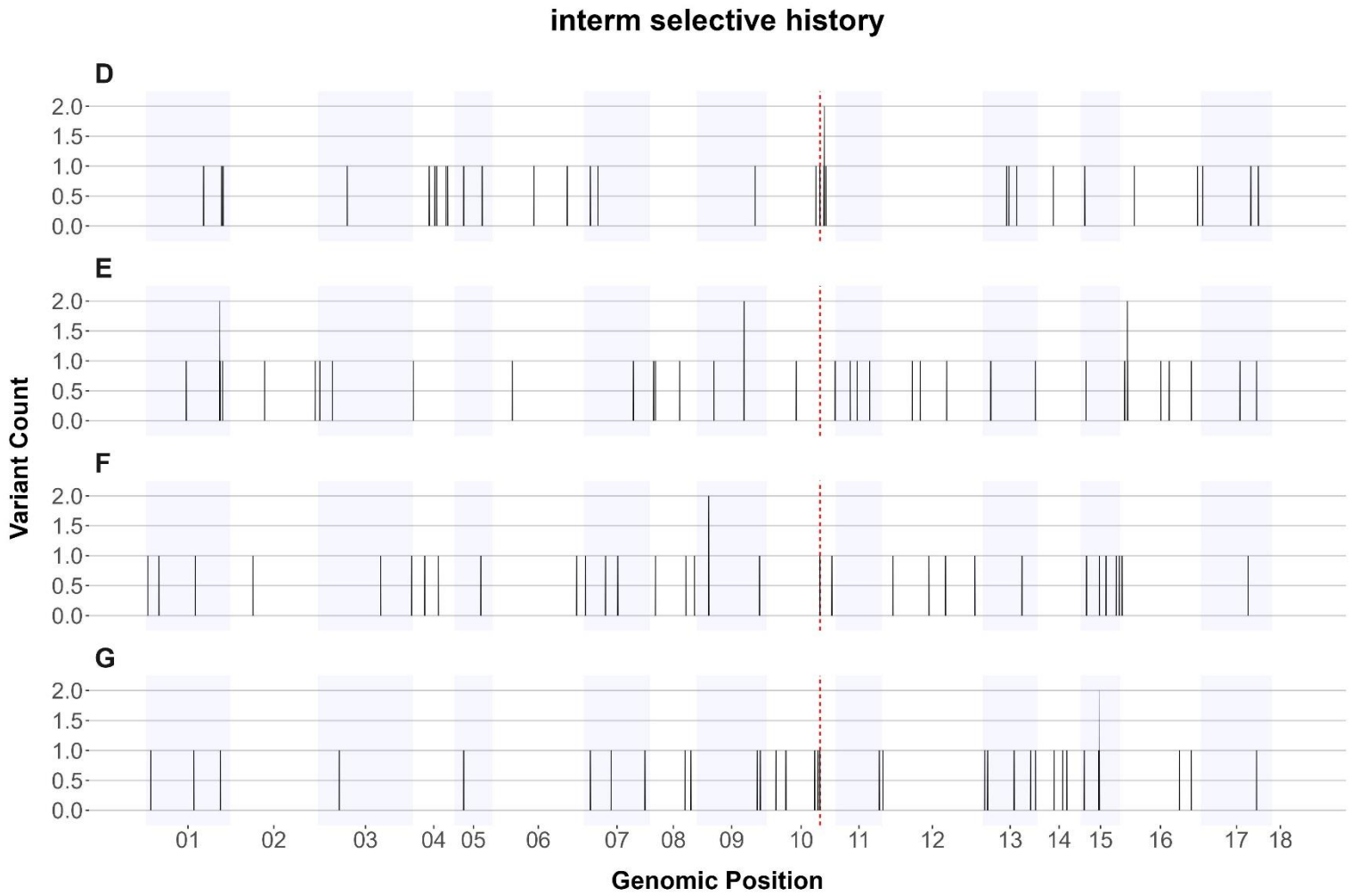

Figure 28: Genome distribution of variants in each of the replicates from the intermediate selective histories of the VR selection experiment. Variants were counted in overlapping bins (100 Kb sliding window with a 10 Kb step) along the genome. Shaded areas delimit chromosome boundaries. Vertical black lines represent bins with variants and their thickness reflects the number of consecutive bins in a given genomic region while their height reflect the number of variants per bin (or cluster of bins). Vertical red dashed lines represent regions with bins containing variants in all three replicates. Their thickness reflects the number of common bins in the genomic region

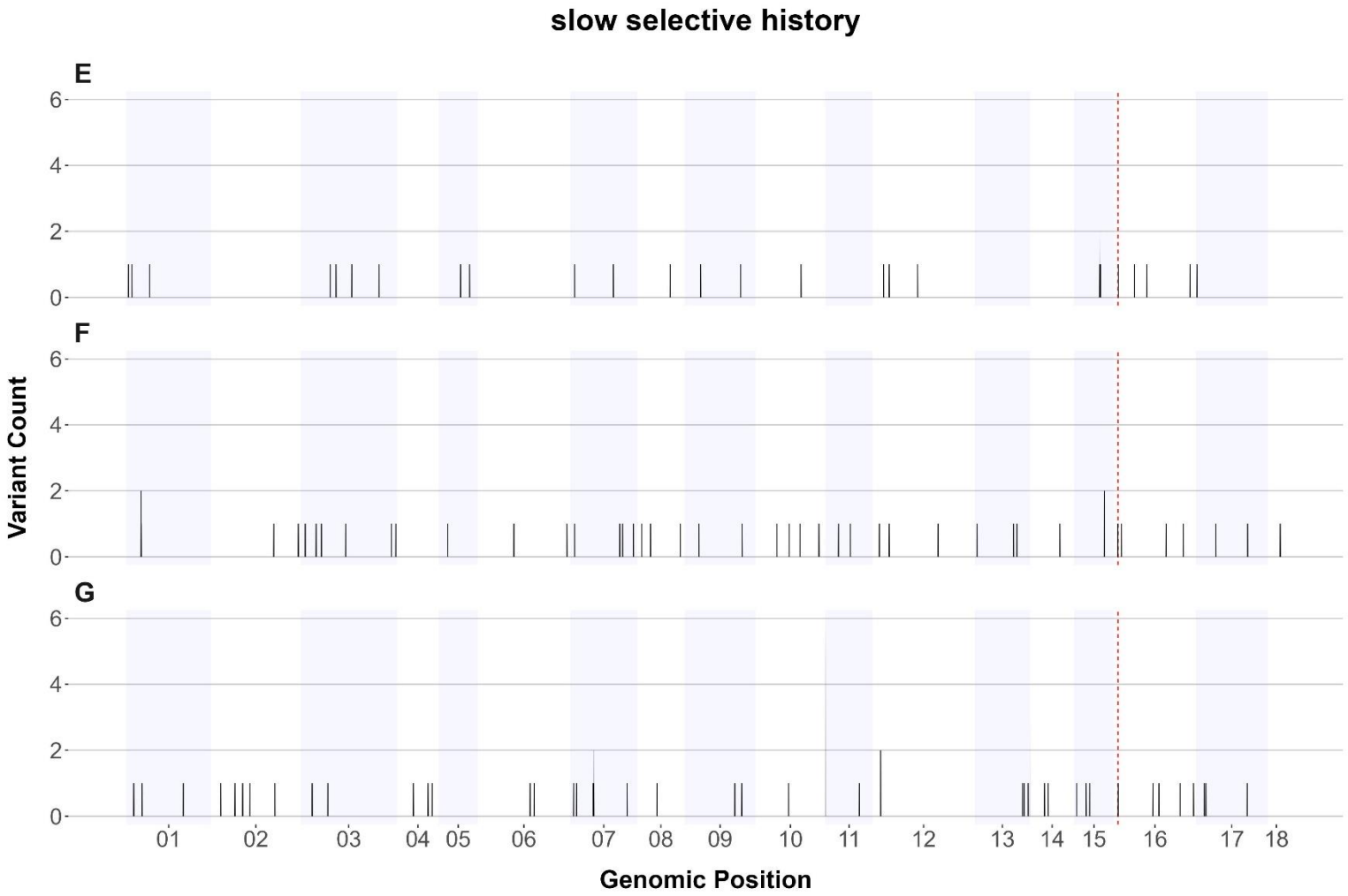

Figure 29: Genome distribution of variants in each of the replicates from the slow selective histories of the VR selection experiment. Variants were counted in overlapping bins (100 Kb sliding window with a 10 Kb step) along the genome. Shaded areas delimit chromosome boundaries. Vertical black lines represent bins with variants and their thickness reflects the number of consecutive bins in a given genomic region while their height reflect the number of variants per bin (or cluster of bins). Vertical red dashed lines represent regions with bins containing variants in all three replicates. Their thickness reflects the number of common bins in the genomic region

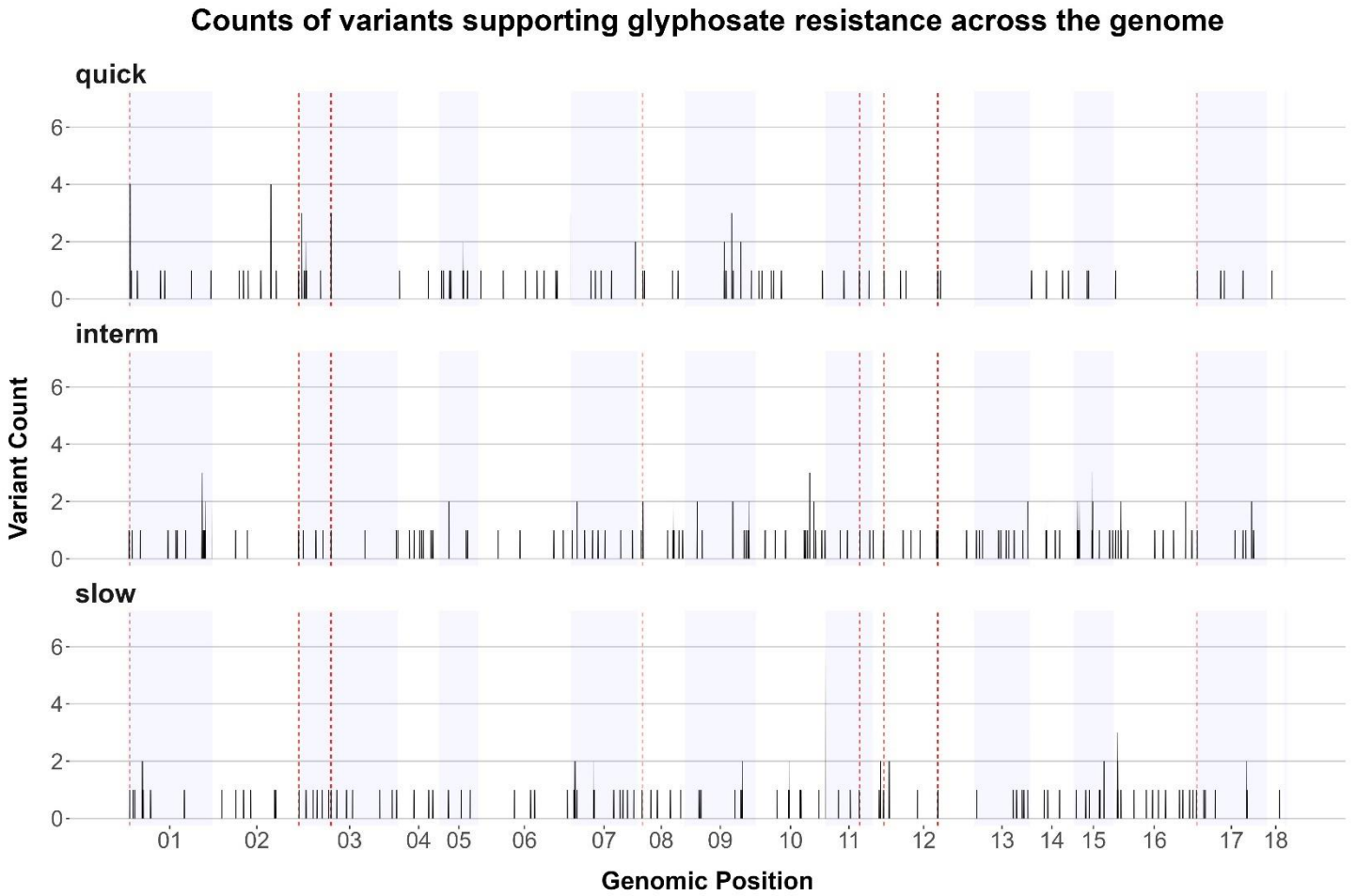

Figure 30: Genome distribution of variants from the quick, intermediate and slow selective histories (all replicates combined) of the VR selection experiment. Variants were counted in overlapping bins (100 Kb sliding window with a 10 Kb step) along the genome. Shaded areas delimit chromosome boundaries. Vertical black lines represent bins with variants and their thickness reflects the number of consecutive bins in a given genomic region while their height reflect the number of variants per bin (or cluster of bins). Vertical red dashed lines represent regions with bins containing variants in all three replicates. Their thickness reflects the number of common bins in the genomic region.

When comparing the effect of selective histories on variants putatively linked to glyphosate resistance, several genomic regions located on chromosomes 1,3, 8, 11, 12 and 17 (Figure 30) contained variants from the list that are common to all selective histories.

5.3.5 Commonalities of variants between replicates

SnpEff annotation of the variant lists based on potential impact returned gene models that may be affected by variants potentially linked to glyphosate resistance. These highlighted gene names were used as a basis for comparisons between replicates in all selective histories from the VR experiment (Figure 31, Figure 32 and Figure 33).

The quick selective history (Figure 31) returned 200 gene names, five of which were consistent in all four replicates, while none of the 285 gene names returned by the intermediate selective histories (Figure 32) or the 182 gene names returned by the slow selective histories (Figure 33) were shared. Functional annotation of the five gene models with variants putatively linked to glyphosate resistance that are common to replicates in the quick selective history (Figure 31) were collated (Table 22).

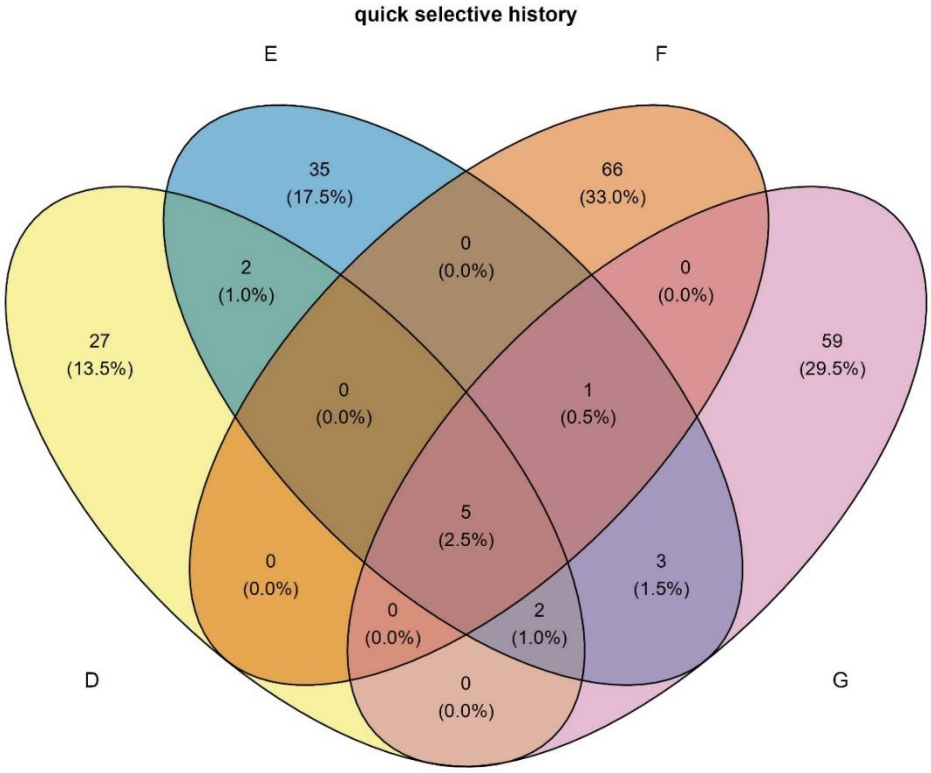

Figure 31: Number of gene names associated with variants putatively linked to resistance in the samples with the quick selective history in the VR selection experiment. Gene names from each replicates are colour coded (yellow: replicate D, blue: replicate E, orange: F and pink: replicate G). Annotation via SnpEff (v4.3+T.galaxy2).

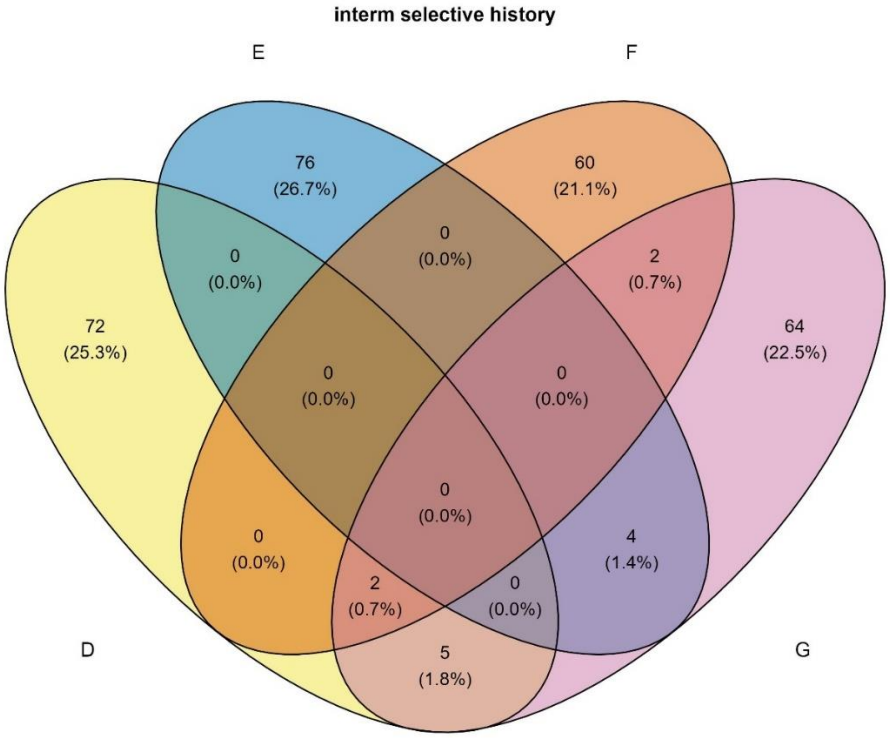

Figure 32: Number of gene names associated with variants putatively linked to glyphosate resistance in the samples with the intermediate selective history in the VR selection experiment.Gene names from each replicates are colour coded (yellow: replicate D, blue: replicate E, orange: F and pink: replicate G). Annotation via SnpEff (v4.3+T.galaxy2).

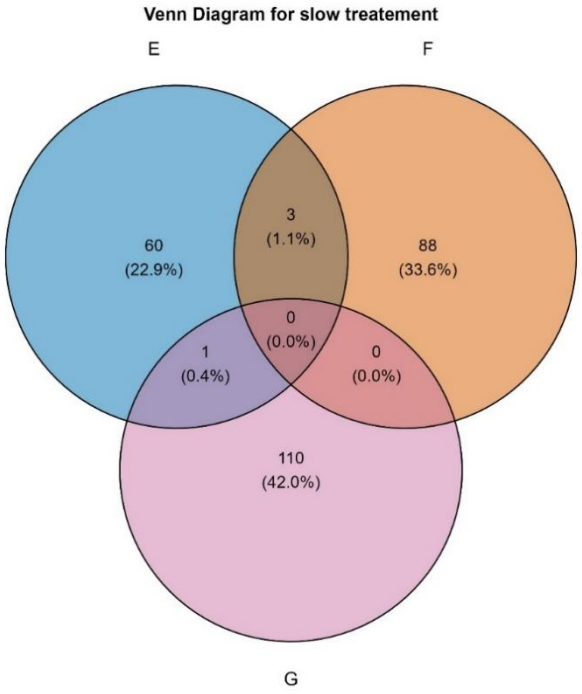

Figure 33: Number of gene names associated with variants putatively linked to glyphosate resistance in the samples with the slow selective history in the VR selection experiment.Gene names from each replicates are colour coded (blue: replicate E, orange: F and pink: replicate G). Annotation via SnpEff (v4.3+T.galaxy2).

Table 22: Genes associated with variants putatively linked to glyphosate resistance that were common to the four replicates of the quick selective history from the VR experiment. Annotation via SnpEff (v4.3+T.galaxy2) Description and Associated PlantFAMS information from Phytozome (https://phytozome-next.jgi.doe.gov/).

Gene names	Description	Associated PLantFAMS	
		Viridiplantae	Chlorophyte
Cre01.g001678	PTHR23257//PTHR23257:SF474 - SERINE-THREONINE PROTEIN KINASE	Tyrosine kinase specific for activated (GTP-bound) p21cdc42Hs	Tyrosine kinase specific for activated (GTP-bound) p21cdc42Hs
Cre01.g001685	NA	PTHR15535//PTHR15535:SF23 - TRANSMEMBRANE PROTEIN 2-RELATED	PTHR15535//PTHR15535:SF23 - TRANSMEMBRANE PROTEIN 2-RELATED
Cre02.g116750 (ATP1A)	Mitochondrial F1F0 ATP synthase, alpha subunit	ATP synthase, putative, expressed	F0F1-type ATP synthase, alpha subunit
Cre02.g116800	PTHR12381:SF46 - SAP DNA-BINDING DOMAIN-CONTAINING PROTEIN	SPRY-domain containing protein, putative, expressed	Serine/threonine protein kinase
Cre02.g116850 (HLM6)	Histone-lysine N-methyltransferase	zinc finger protein, putative, expressed	1.14.11.4 - Procollagen-lysine 5-dioxygenase / Procollagen-lysine,2-oxoglutarate 5-dioxygenase

5.3.6 Gene Ontology analysis

Gene Ontology (GO) terms for molecular function (MF), cellular components (CC) and biological processes (BP) were extracted for annotated genes models from the list of putative variants supporting glyphosate resistance comprised of both the SR and VR selection experiments (Figure 34, Figure 35 and Figure 36). There were notable rare consistencies: ATP binding and protein kinase MF (Figure 34), Membrane CC (Figure 35) and protein phosphorylation BP (Figure 36) were found across the SR experiment and the quick and intermediate selective histories of the VR experiment.

Figure 34: Molecular function GO terms heatmap for both SR and VR experiments.(N.B this figure is available in a PDF format for improved visualisation, see Appendix 7).

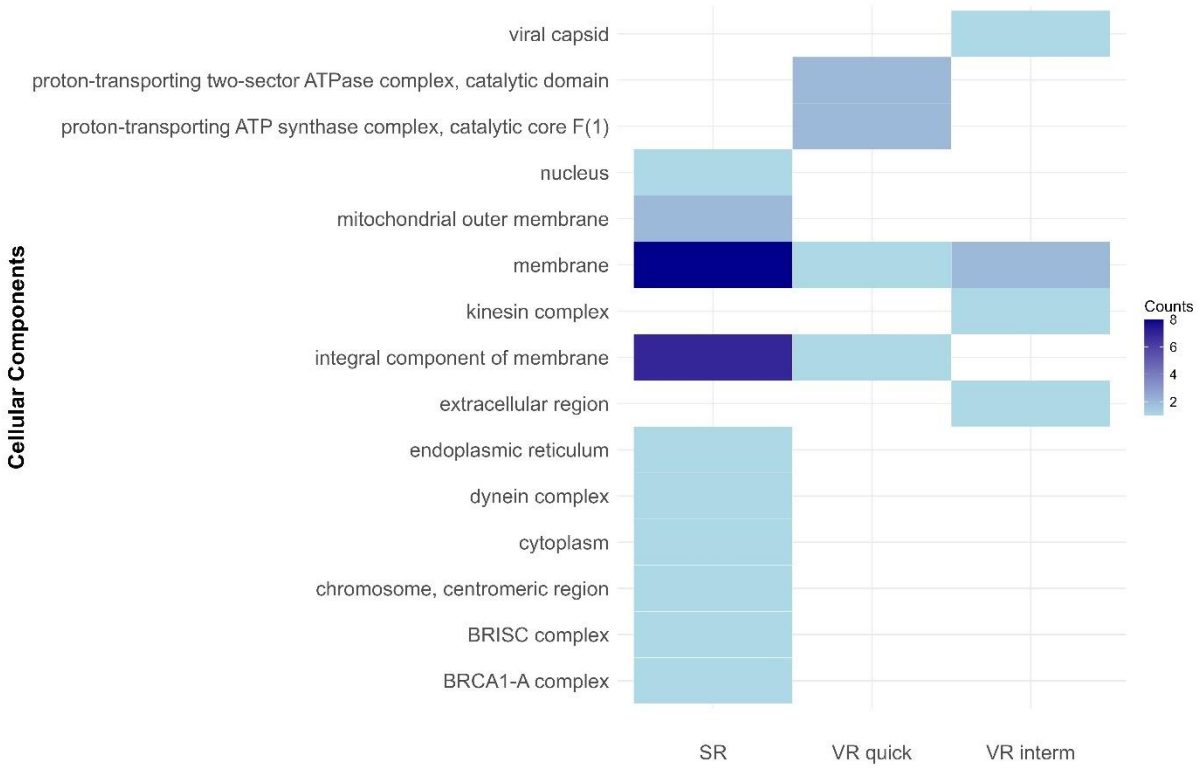

Figure 35: Cellular component GO terms heatmap for both SR and VR experiments.(VR slow selective history gene names with associated GO terms did not have cellular component descriptions

.

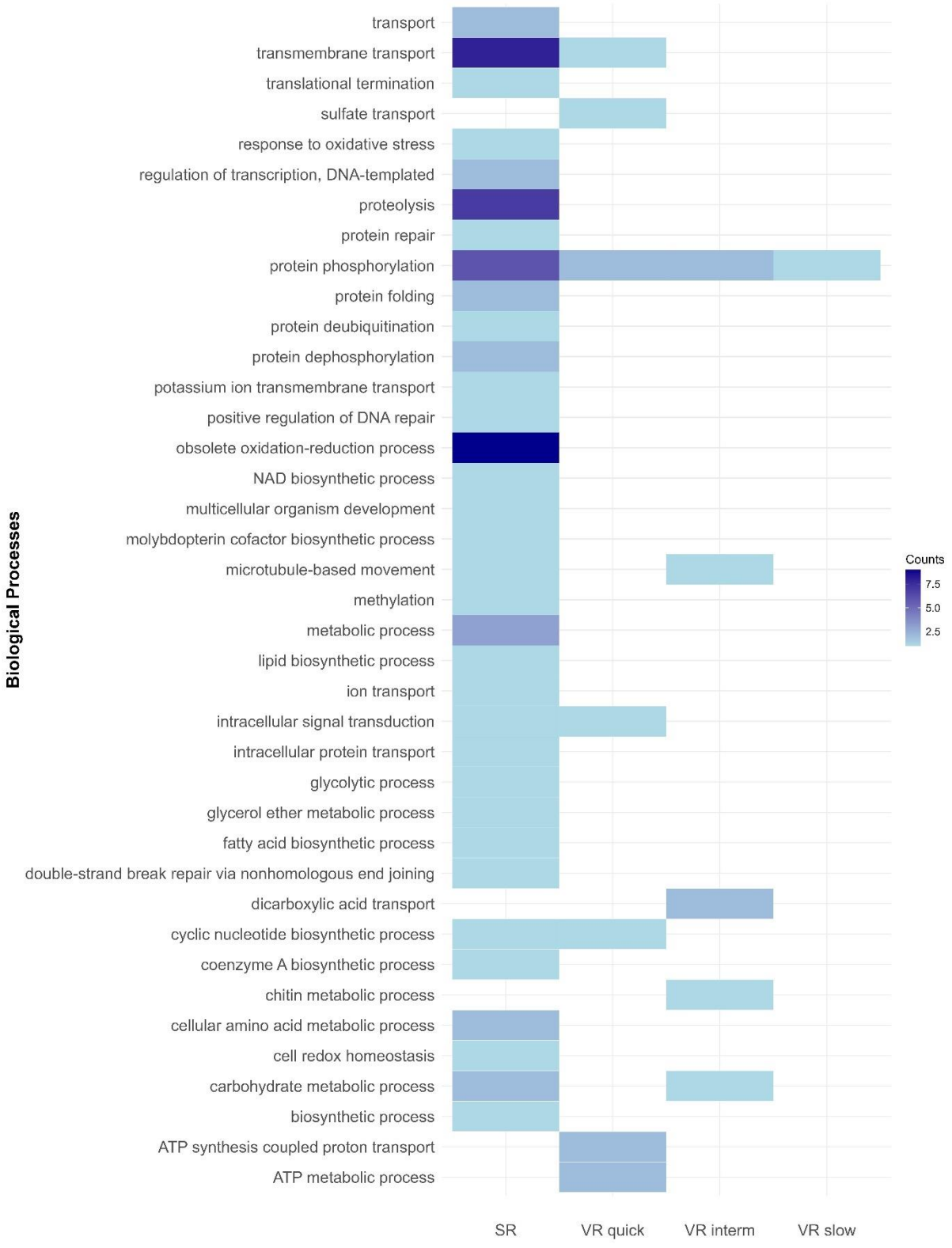

Figure 36: Biological processes GO terms heatmap for both SR and VR experiments.

.

5.4 Discussion

I investigated the impact of different selective histories generated via selection with contrasting treatments in the VR experiment (presented in Chapter 4) and I attempted to link the genetic changes to the observed glyphosate resistant phenotype. Glyphosate treatments applied in the single-rate (SR) experiment within Chapter 3, and the VR-quick here were identical (exposure to glyphosate at 1 MIC). Consequently, results from these two experiments can be compared, albeit bearing in mind that clonality of the genomes at the onset of selection can be assumed for all samples in VR but not in SR.

The analysis presented in this chapter is based on variant calling and filtering methods developed to link the genotype to the phenotype in the SR experiment (presented in Chapter 3). Hence, similarly to Chapter 3, analysis within this chapter involves investigation of the list of putative variants linked to glyphosate resistance focussing on small variants (already discussed in Chapter 3). Limitations of the variant calling analysis and the filtering methods applied to obtain the final list of putative variants were also previously discussed in Chapter 3.

The presence of mutation(s) in *EPSPS* was investigated and no evidence was found for the presence of glyphosate resistance endowing *EPSPS* mutations (which would have been indicative of glyphosate TSR mechanism) in any of the selective histories. Also, a preliminary informal investigation of read depth around the *EPSPS* loci did not suggest the presence of *EPSPS* copy number variation (CNV) in the GR samples. Absence of variants supporting glyphosate resistance in *EPSPS* and absence of *EPSPS* CNV was consistent with findings in the SR experiment (Chapter 3).

These results suggest that, in both studies, if glyphosate TSR is involved, at the genome level, it would have to be through mutation(s) in a distant genomic region(s) affecting *EPSPS* expression, or through epigenetic modification. Further analysis of the existing dataset with STRING (Search Tool for the Retrieval of Interacting Genes/Protein) could be performed to identify genes functionally connected with EPSPS and investigate if they participate in shared pathways. Further work could include genome-wide structural variation (SVs) analysis, allowing investigation of the potential impact of such variants on the observed GR phenotype and to further confirm the absence of such variant affecting the *EPSPS*. TSR and NTSR mutations can occur independently or concurrently in individual plants (Alcántara-de la Cruz et al., 2016; Larran et al., 2022), resulting in complex resistance profiles. Therefore, mutations underpinning glyphosate NTSR mechanisms are likely to be present in the final list of variants. This suggests that in both studies, the observed *C. reinhardtii* glyphosate resistant phenotype is likely supported by NTSR mechanisms. This was observed in all replicates exposed to different treatments (SR and VR) suggesting that the results are independent of the selective histories tested.

The final list of variants potentially contains variants underpinning glyphosate NTSR. NTSR mechanisms are more challenging to characterise due to their diversity and polygenic nature (Délye and Christophe, 2013). In this thesis chapter, the final variant list has been explored to assess likelihood that variants support GR phenotype through NTSR mechanisms with simple diagnostic tools to narrow down to potential important groups of variants to be prioritised for further study and functional validation of their impact on the observed GR phenotype. Further work could use

methods such as CRISPR/Cas9 gene editing (Ghribi et al., 2020), RNA interference (Cerutti et al., 2011), and gene overexpression (Hema et al., 2007) for functional validation of gene function in *C. reinhardtii*. Additionally, comparisons are made between selective histories to identify the effect of rate of change in glyphosate exposure in the number and type of variants observed.

First, SnpEff variant categories and predictions were used to investigate the final list of variants. Regarding variant categories, results were similar to those of the SR experiment: insertions constituted the majority of variant types. These results are contrasting with the findings of Ness et al. (2015) but consistent with evidence that stress may increase the rate of indel mutations in *C. reinhardtii* (Hasan et al., 2022). The possibility that some of these insertions are false positives was discussed in Chapter 3.

Concerning SnpEff annotations, few variants were annotated by SnpEff as being located in a coding region in all selective histories. This suggests that variants linked to the observed glyphosate resistant phenotype could be mainly found in regions impacting gene expression (such as UTRs, enhancers, silencers, transcription factor binding sites, promoter regions etc..). This is consistent with the fact that only seven variants were classified 'HIGH' impact on gene function, through disrupting gene's reading frame (frameshift variants). Similarly to results observed in the SR experiment, these results suggest that putative variants linked to a GR phenotype might not lead to amino acid changes in encoded proteins, but might instead affect GR by other means, such as altered gene expression of enzymes involved in glyphosate metabolism or detoxification. Gene models predicted to be affected by HIGH impact variants were found for the quick, intermediate and slow selective histories.

The major limitations of the results presented above are linked to the predictive nature of SnpEff annotations. Although informative and potentially useful to help decide which candidate genes to prioritise for further investigation, the effect of variants on the genotype can only be demonstrated with functional validation of the gene function. Additionally, these predictions rely on the quality of the available reference genome assembly and annotations.

As an additional screening tool, genomic variant distribution was investigated for the presence of clusters signalling any obvious "glyphosate resistance hotspots". Consistent with results of the SR experiment, there were no obvious clusters of variants, providing no evidence for glyphosate resistance hot spots in any of the three selective histories (Figure 27, Figure 28 and Figure 29).

The presented data shows little overlap between the replicates in the quick (on chromosome 3 and 9, Figure 27), intermediate (chromosome 10, Figure 28) and slow (chromosome 16, Figure 29) selective histories. These results suggest that after 12 weeks, although there is evidence of convergent evolution at the phenotypic level (i.e. all populations evolved resistance to glyphosate), there is no strong evidence of convergent evolution under relatively fast rates of change at the genomic level. Differences at the genomic level may explain the differences in observed fitness cost of glyphosate resistance between the *C. reinhardtii* populations from the quick and intermediate selective histories (Figure 17). There is evidence that the rate of environmental change (selective history) has an impact on fitness cost of resistance (Collins, 2011; Collins and De Meaux, 2009; Collins et al., 2007) and previous studies have linked glyphosate resistant endowing mutations to fitness cost

in weeds (Vila-Aiub et al., 2019). Should further work lead to a list of prioritised variants for functional validation, both their impact on the GR phenotype and their impact on fitness cost should be investigated.

A limitation to the interpretation of these results is linked to the fact that selective histories had different impacts on *C. reinhardtii's* population growth during the VR selection (Chapter 4) inherently leading to different number of generations being exposed to contrasting conditions. Over the course of 12 weeks, GS populations in non-selective environments might be expected to have evolved over approximately 120 generations, unlike GR populations exposed to quick rates of change (Figure 16).Although not directly measured, it can be assumed that GR populations exposed to the intermediate and slow selection histories underwent approximately 120 generations, based on their evolutionary dynamics resembling those of the control (Figure 16). A better approach would be to express time in generations which is more informative because evolutionary changes depend on the number of reproductive cycles rather than absolute chronological time. This would allow more robust conclusions on the effect of selective history on level of convergence.

Comparing variant distribution between the quick, intermediate and slow selective histories highlighted several 100 Kb genomic regions harbouring putative variants linked to GR, with some common to all selective histories (Figure 30). Although not shared by all replicates in each selective history, these genomic regions may potentially harbour key genetic variants responsible for glyphosate resistance, selected across both environments. It is important to note, however, that there is a chance these locations simply signify problematic regions of the genome more prone to false positives. Further work narrowing down into these genomic regions is required to conclude if any chromosome harbours a glyphosate resistance hot spot. If such region does not seem more prone to false positive than the rest of the genome, investigating if genes products or regulatory element with variants in these regions take part in known glyphosate resistance metabolic pathways. This was not achieved during this thesis due to time constraints.

Further analyses were undertaken to examine whether variants that are specific to each selective history could reveal glyphosate resistance mechanisms specifically arising under a given selective pressure. Sets of gene models that may be impacted by variants from the final list according to SnpEff annotations were compared between replicates of each of the quick, intermediate and slow selective histories (Figure 31, Figure 32 and Figure 33).The large number of genes unique to each replicate suggests a significant level of variability in the genetic response to glyphosate in individuals from the quick (Figure 31, there were only five overlapping genes), intermediate (Figure 32, no overlapping genes) and slow (Figure 33, no overlapping genes) selective histories. These results reveal distinct patterns of genetic adaptation in the under sudden changes in doses (quick selective history) and gradual changes in doses (intermediate and slow selective histories). Selection with quick rates of glyphosate dose increase fosters partial convergence with high individual variability while selection with intermediate or slow rates of glyphosate dose increase leads to heterogeneous responses with no shared genes. This may suggest that, under stronger selective pressure, a selection of few variants could be essential to glyphosate adaptation. These findings highlight the interplay between selection tempo, evolutionary dynamics, and the predictability of adaptive outcomes(Collins and De Meaux, 2009; Collins et al., 2007; Perron et al., 2006, 2008). In this list of genes (Table 22), two are regulatory genes (*Cre01.g001678* and *Cre02.g116850*) and one is metabolism-related (*Cre02.g116750)*. This suggest that under sudden increase in glyphosate dose,

glyphosate resistance is driven by a complex, NTSR and polygenic mechanism(s) and might provide an explanation for the observed diverse evolutionary pathways taken by different replicates. Should this be confirmed by further work, it would underscore the importance of monitoring NTSR glyphosate resistance mechanisms.

Considering that there are still suspected false positive variants in the final list which could not be refined (as discussed in in Chapter 3) due to time constraints,

the number of variants in the final list could not be narrowed down further to fewer candidate variants to take forward for functional validation of their effect on glyphosate resistance.

To evaluate the likelihood of returned variants being involved in GR mechanisms, I selected gene models with predicted HIGH impact putative variants linked to GR and the gene models with putative variants linked to glyphosate resistance in the quick selective history. Table 23 summarises information available on putative "candidate gene" functions and the potential evidence of roles in stress tolerance. One particularly noteworthy gene is *Cre09.g400950,* classified as belonging to the major facilitator superfamily (MFS) and predicted to be affected by high impact variants in replicate F from the intermediate selective history (Table 21). There is evidence suggesting that MFS proteins are involved in NTSR glyphosate resistance mechanisms in bacteria (Staub et al., 2012; Tao et al., 2017).MFS proteins are integral membrane transporters shown to play a role in plants response to stress (Haydon and Cobbett, 2007; Remy et al., 2013). Studies showed that MFS membrane transporters confer resistance to fungicides (Ramon-Carbonell et al., 2019) and antibiotics (Wan et al., 2023) through increased efflux (Ramon-Carbonell et al., 2019 ; Atin et al., 2017). Therefore it would be interesting to investigate if *Cre09.g400950* effectively codes for an MFS transporter that extrudes glyphosate from *C. reinhardtii* cells thus conferring resistance. Although functional validation is necessary to ascertain the role of these variants in supporting the glyphosate resistant genotype, such results are preliminary evidence that existing stress response pathways in *C reinhardtii* may be implicated here in the observed evolutionary response to glyphosate. Furthermore, it indicates that the method developed in this chapter could be successful in linking relevant variants to the observed glyphosate resistant phenotype provided further improvements are implemented.

In future work the use of InterProScan could help predict gene function by identifying conserved domains, motifs, and active sites, integrating data from multiple protein signature databases, and providing standardized functional annotations like GO terms and pathway associations (Blum et al., 2025; Jones et al., 2014).

Table 23: Subset of putative "Candidate genes" from the VR experiment with described gene function . Gene information from SnpEff (v4.3+T.galaxy2). Descriptions of putative gene product from various sources (Phytozome-CC503_V5.6 annotations or PlantFAMS) was retrieved from JGI's Phytozome13 Portal (https://phytozome-next.jgi.doe.gov/). Comments summarise potential links to stress tolerance from peer-reviewed sources.

Description	Description source	Gene	Comment
16S rRNA (cytosine(967)-C(5))-methyltransferase	(Phytozome V5.6 annotation)	*Cre16.g647602*	Cytosine methyltransferases are enzymes responsible for adding methyl groups to cytosine residues in DNA or RNA, affecting gene expression and other cellular processes. There is evidence of methyltransferases role in biotic and abiotic stress tolerance (Bvindi et al., 2022).
Histone-lysine N-methyltransferase	(Phytozome V5.6 annotation)	*Cre02.g116850* (HLM6)	Histone-lysine N-methyltransferases play significant roles in modulating stress responses by altering chromatin structure and gene expression. There is evidence for their role in biotic and abiotic stress in tomato (Bvindi et al., 2022).
major facilitator superfamily	(Phytozome V5.6 annotation)	*Cre09.g400950* (NCT2)	MFS transporters in plants are integral to various physiological processes, including hormone transport, metal ion homeostasis, and responses to both abiotic and biotic stresses in plants and microorganisms. There is evidence suggesting MFS proteins are involved in glyphosate resistance in bacteria (Staub et al., 2012; Tao et al., 2017).
Mitochondrial F1F0 ATP synthase, alpha subunit	(Phytozome V5.6 annotation)	*Cre02.g116750* (ATP1A)	F_1F_0-ATP synthase is integral to energy production and stress tolerance mechanisms. Its activity is essential for maintaining cellular energy homeostasis under stress conditions. There is evidence for a role in heat stress (Liu et al., 2021) and salt stress (Soontharapirakkul et al., 2011) tolerance.
psaA mRNA trans-splicing factor	(Phytozome V5.6 annotation)	*Cre10.g440000* (OPR120,RAA8)	Trans-splicing factors involved in *psaA* mRNA maturation are crucial for the proper assembly of photosystem I in chloroplasts (Lefebvre-Legendre et al., 2015).
PTHR23033 - BETA1,3-GALACTOSYLTRANSFERASE	(PlantFAMS-Chlorophytes& Viridiplantae)	*Cre11.g467750*	β-1,3-Galactosyltransferases (GalTs) are enzymes that play a crucial role in the biosynthesis of arabinogalactan proteins (AGPs), which are important components of the plant cell wall. AGPs are involved

			in various plant growth and development processes, including cell expansion, somatic embryogenesis, and responses to environmental stresses (Zhang et al., 2020; Zhao et al., 2019). There is evidence that AGPs (C. Zhao etal., 2019) and GalTs (Li et al., 2013) play a role in salt stress tolerance in Arabidopsis.
Serine/threonine protein kinase	(Phytozome V5.6 annotation)	*Cre01.g001678*	Serine/threonine protein kinases are integral to the stress tolerance mechanisms of both plants and microbes. By modulating various signalling pathways, these kinases enable organisms to adapt to and survive under diverse stress conditions. There is evidence for their role in drought (Lim et al., 2020) and salt stress tolerance (Zhang et al., 2019) in plants.
	(PlantFAMS-Chlorophytes)	*Cre02.g116800*	
TMEM2- related	(PlantFAMS-Chlorophytes& Viridiplantae)	*Cre01.g001685*	TMEM2 (in plants homologs are named TMEM2-related proteins) is a trans membrane protein that degrades and regulates levels and function of hyaluronan (aka hyaluronic acid HA) (Yamamoto et al., 2017). HA is present in the extracellular matrix (ECM) and plays a crucial role in maintaining cell wall integrity, cell signalling facilitation and modulating responses to environmental stresses. There is evidence that HA plays a multifaceted role in stress tolerance by modulating inflammation(Petrey and de la Motte, 2014), protecting against oxidative damage and facilitating tissue remodelling (Berdiaki et al., 2023) in humans.

The use of Gene ontology (GO) terms was also used to narrow down the list of variants to prioritise for further analysis. GO term are standardised descriptions of gene products' molecular function (MF), localisation (cellular component -CC) and biological processes (BP) in which it may be involved. GO terms were extracted from annotated gene models from the list of putative variants supporting glyphosate resistance, comprised of both the SR and VR selection experiments. GO term enrichment analysis was attempted but remain inconclusive due small sample size leading to lack of statistical power. Consequently, heatmaps were generated to compare lists of GO terms associated with gene models observed in the SR and VR experiment (Figure 34,Figure 35 and Figure 36).

Given that SR and VR-quick samples were subjected to the same glyphosate treatment (constant exposure to glyphosate at 1 MIC), one might have expected their results to have looked more similar. However, the SR and VR selection experiments lasted less than 200 generations in the controls (and thus even fewer generations for the selected samples) which might not be enough to observe convergence under such selection pressure. The main bias in the representation of the results is that genes without annotations and associated GO terms are excluded.

There are rare and noteworthy consistencies between the SR, VR quick,VR intermediate and VR slow selective histories: ATP binding and protein kinase MF (Figure 34), Membrane CC (Figure 35) and protein phosphorylation BP (Figure 36) are found across SR and all VR selective histories. Their consistency may suggest that the protein phosphorylation pathway, membranes and ATP binding molecular functions as well as protein kinases, could represent a signal for glyphosate resistance. Although GO terms provide useful insights, they represent broad groupings of MF, CC and BP, thus specific components of these ontologies responsible for GR still need further investigation. Some MF GO terms associated to genes annotated as affected by variants in final list suggest putative variants may support glyphosate resistance phenotype as enzymes with similar function were associated with glyphosate resistance (Table 24).

Table 24: Molecular function (MF) Gene Ontology (GO) terms annotations of gene annotated as impacting genes from gene families with known role in glyphosate resistance in other organisms. Annotations from (Phytozome V5.6 annotation).

GO term	Comments
Oxidoreductase activities	Oxidoreductases (in fungi) and glyphosate oxidoreductase (GOX -in bacteria) may play a role in glyphosate detoxification (Firdous et al., 2020; Guo et al., 2021; Zhao et al., 2015).
Monooxygenase activity	Monooxygenases are involved in phase I detoxification pathways. CP450 are monooxygenases and have been suspected to participate in glyphosate detoxification in plants (Deng et al., 2022; Van Etten et al., 2020; Laforest et al., 2020; Piasecki et al., 2019a), and in fungi (Mesnage et al., 2020).
Acetyl transferase activity	Glyphosate acetyltransferase (GAT) catalyses glyphosate N acetylation in bacteria (Castle et al., 2004; Shushkova et al., 2016).
Transporter activity	Several membrane transporters were identified as playing a role in glyphosate resistance : GltP and GltT (Wicke et al., 2019), Pdr5 (Ravishankar et al., 2020b), ABC transporters in microbes and in plants (Deng et al., 2022; Gerakari et al., 2022; Laforest et al., 2020; Moretti et al., 2017; Pan et al., 2021; Peng et al., 2010; Piasecki, Carvalho, et al., 2019; Piasecki, Yang, et al., 2019;Schneider et al., 2021; Tani et al., 2015; Van Etten et al., 2020) (See Chapter1 Table 1 for summary)
Hydrolase activity	Hydrolyse encoding gene was differentially expressed gene in glyphosate tolerant fungi (Guo et al., 2021).

The interpretation of this data requires caution since it relies on the quality of the available genome for the variant calling results and the quality of the available annotation for gene models and GO term assignment to variants. In other words, it is possible that the final list of variants still contains false positives (FPs) and that variants effectively supporting the GR phenotype impact unknown gene or genes without functional annotations. This analysis could be improved with additional samples sequenced to provide enough statistical power for GO term enrichment analysis, although this would require duplicating the VR experiment.

In conclusion, the work presented in this chapter build upon findings described in Chapter 3 and Chapter 4 to investigate the genomic basis of glyphosate resistance in *C. reinhardtii* subjected to different rates of glyphosate dose escalation: quick (1 MIC), intermediate (incremental dose increases up to 1 MIC in three steps) and slow (incremental dose increase up to 1 MIC in six steps). Acknowledging the methodological limitations of this study, the findings align largely with those reported in Chapter 3, where *C. reinhardtii* populations adapted to glyphosate at 1 MIC. Specifically, no evidence was found for variants associated with glyphosate resistance in the *EPSPS* gene or for *EPSPS* copy number variation (CNV). None of the variant in the final list were predicted to be located in coding regions, nor was there evidence of a clear "glyphosate resistance hotspot" in the *C. reinhardtii* genome. Analysis of a reduced set of genes annotated as impacted by variants, along with Gene Ontology (GO) terms associated with all variants in the final list, suggests that glyphosate resistance is at least partially supported by pre-existing stress response pathways, indicating a potential role for non-target-site resistance (NTSR) mechanisms. Since these results are consistent across both single rate (SR) and variable rate (VR) experiments, it appears that the glyphosate selective history has minimal influence on these outcomes under the conditions tested in this thesis.

However, selective history appears to influence the distribution of genomic variants and the degree of convergence across replicates. Specifically, partial overlap in gene models affected by variants was observed under rapid selective conditions, whereas greater heterogeneity was noted under intermediate and slow selective conditions. This suggest that under the harsher selective history, a set of few variants might be instrumental to glyphosate adaptation in *C. reinhardtii.*

6 General discussion

6.1 Brief summary of findings

Intensive and prolonged glyphosate use has imposed strong selective pressures, driving the evolution of resistance in weeds and contaminating non-target ecosystems. Understanding the evolutionary dynamics and outcomes of glyphosate selection is critical for addressing these challenges. The evolution of glyphosate resistance follows diverse pathways, as evidenced by the variety of resistance mechanisms observed. In higher plants, glyphosate resistance is supported by both target-site resistance (TSR) and non-target-site resistance (NTSR) mechanisms. TSR mechanisms include mutations in the EPSPS-encoding gene, increased gene copy number, enhanced transcription, or elevated EPSPS activity (Galeano et al., 2016; Jander et al., 2003; Koo et al., 2018; Molin et al., 2017; Zhang et al., 2015). Meanwhile, NTSR mechanisms identified thus far include reduced glyphosate uptake, decreased translocation, enhanced metabolic degradation, and detoxification (Deng et al., 2022; Michitte et al., 2007; Pan et al., 2019; Vila-Aiub et al., 2011). Previous research has shown that *Chlamydomonas reinhardtii* can evolve resistance to glyphosate, with studies exploring the evolutionary dynamics and phenotypic outcomes of this process (Hansson et al., 2024; Lagator et al., 2012, 2013, 2014a; Melero-Jiménez et al., 2021; Vogwill et al., 2012). However, no study has yet examined the connection between the evolution of a glyphosate-resistant phenotype and associated genetic changes.

This thesis explores the evolution of glyphosate resistance in *C. reinhardtii* subjected to different selection regimes and examines their effects on both the phenotype and associated genomic changes using an evolve and re-sequence (E&R) approach.

First, I aimed to develop a method for linking observed phenotypic changes to underlying genomic alterations starting with a simple experimental design (*C. reinhardtii* lines exposed to a single rate (SR experiment) of change in glyphosate dose (1 MIC) (Chapter 3). In a second selection experiment, *C. reinhardtii* populations were subjected to variable rates (VR experiment) of glyphosate dose escalation (quick, intermediate, and slow). This allowed an investigation into the impact of different rates of environmental deterioration on adaptation at the phenotypic level by assessing the influence of selective history on the evolution of resistance and the associated fitness costs (Chapter4). Finally, examining the genomic basis of glyphosate resistance in *C. reinhardtii* exposed to three distinct rates of glyphosate dose escalation: quick selective history (1 MIC), intermediate selective history (three incremental dose increases up to 1 MIC) and slow selective history (six incremental dose increases up to 1 MIC) was achieved building up on work from the previous chapters (Chapter 5).

6.1.1 Phenotypic and fitness adaptation in response to glyphosate selection

The impact of glyphosate selection on phenotype was assessed looking at the evolution of resistance (SR and VR) and associated fitness cost (VR only). Fairly consistent phenotypic evolution of glyphosate resistance was observed under selection at doses up to 1 MIC (except for one replicate in the VR selection experiment).This result is consistent with previous work (Hansson et al., 2024; Lagator, 2012c; Lagator et al., 2014a, 2014b; Melero-Jiménez et al., 2021) and confirm that resistance evolution is a consistent outcome of glyphosate selection at 1 MIC in *C. reinhardtii.*

Selective history impacted the evolutionary dynamic of glyphosate adaptation with the quick selective history (rapid rate of change) reducing population size, increasing variability in the evolutionary dynamic and delaying evolution of resistance. The quick selective history drastically impacted population size at the onset of selection and immigration was required to maintain population and avoid extinction until the occurrence of evolutionary rescue. This is consistent with results from experiments using a weekly transfer into fresh media following a source-sink scenario (Lagator, 2012c; Lagator et al., 2014a, 2014b). In contrast, a more gradual change in dose, as imposed by intermediate and slow selective histories, did not impact population size. Consistent with previous work in this species (Lagator, 2012; Lagator, et al., 2014a; Lagator, et al., 2014b; Melero-Jiménez et al., 2021),when resistance evolved, it was heritable. This confirms glyphosate adaptation occurred through mutations(s) and or heritable epigenetic modifications.

Regardless of the selective history, *C. reinhardtii* populations did not adapt to doses above 1 MIC. These findings align with Hansson et al. (2024) but not with results reported by Melero-Jiménez et al.(2021) who reported *C. reinhardtii* population adaptation up to 1.8 MIC. I hypothesize that the observed shift from the initial MIC is driven by the rate of glyphosate dose increase. Specifically, a ratchet protocol, which raises the glyphosate dose only after a large population size is achieved (Melero-Jiménez et al., 2021), thus imposing a slower rate of change compared to the conditions tested in other studies.

Selective history also impacted fitness costs associated with glyphosate resistance: the fitness in ancestral environment of populations experiencing intermediate and slow selective histories being lower than in populations experiencing the quick selective history. I hypothesised that higher fitness cost are likely due to accumulation of multiple mutations of smaller effect and their associated fitness cost (Collins et al., 2007; Perron et al., 2008). Similar findings have been reported previously: a slow glyphosate dose increase using a ratchet protocol caused significant fitness costs (Melero-Jiménez et al., 2021), whereas direct exposure to glyphosate at 1 MIC led to resistance with no associated fitness cost, even at intermediate doses, such as 0.5 MIC (Hansson et al., 2024; Lagator et al., 2012c; Vogwill et al., 2012). Studies reported glyphosate resistance with no associated fitness cost nor cross resistance to other herbicides, suggesting that glyphosate resistance may be supported by TSR mechanisms (Lagator et al., 2012; Vogwill et al., 2012). This remains speculation since the mechanism of glyphosate resistance in *C. reinhardtii* is still unknown to date.

The impact of glyphosate selection on genomic changes was then investigated. The SR and VR glyphosate selection experiments generated replicate pairs of (GR) and glyphosate sensitive (GS)

lines with an experimental evolution approach (i.e. under tightly controlled laboratory conditions) for approximately 180 generations in the SR experiment and 120 generations in the VR experiment.

6.1.2 Genomic adaptation underlying observed Glyphosate resistance

Following evolution of resistance, whole genome sequencing of GR and GS individuals was performed. Variant calling (VC) and variant filtering pipelines were developed on the SR GS and GR samples (Chapter 3) to only retain high quality variants susceptible to support the observed GR phenotype. The performance of these pipeline regarding the number of false positives was greatly dependent on the reference genome used, highlighting the need for highly contiguous assembly of a genome as closely related as possible to the samples under investigation. The same pipelines were then applied to the VR samples (Chapter 5). The final lists of putative variants supporting glyphosate resistance were investigated with diagnostic tools to determine the resistance mechanisms at play (TSR and/or NTSR), compare genomic changes between replicates and attempt to produce a smaller list of variants to prioritise for further work (Chapter 3 and 5). In the case of the VR samples, the impact of selective history on the genomic changes was also investigated (Chapter 5).

Consistencies were found in the results of the SR and VR experiments genomic analysis. Investigation of both final variant lists suggested the absence of two major target-site resistance (TSR) mechanisms: *EPSPS* mutations or *EPSPS* copy number variation (CNV) where not evidenced. Measuring changes in *EPSPS* expression and a structural variant (SVs) analysis would have allowed to completely rule out the presence of glyphosate TSR mechanisms but was outside the scope of this study and could be investigated in future works. The variants in the final list were further investigated *in silico* to estimate their role in supporting the observed GR phenotype. Most variants were annotated as located in non-coding regions suggesting they may reside in regions influencing gene expression. This aligns with the observation that few variants were predicted to have a 'HIGH' impact on gene function, likely through frameshift mutations. Therefore, it is likely that the majority of variants in the final lists will not alter the amino acid sequence of encoded proteins.

In both studies, genomic variant distribution did not reveal the presence of clusters, providing no evidence for glyphosate resistance hot spots in *C. reinhardtii*. Distribution of DEGs identified genomic regions linked to resistance to other herbicides in *Amaranthus tuberculatus* (Giacomini et al., 2020). Consequently, gene models predicted to be impacted by the variants in the final lists in the SR and VR samples were investigated. First, genes annotated as impacted by variants classed by SnpEff (see 2.5.2 for more detail as HIGH impact and genes common to all replicates within a selective history were used as a reduced set of genes to retrieve function annotations. Transmembrane Protein 2-related (TMEM2) protein and protein kinases (PKs) are found in both SR and VR reduced gene set. This consistency suggest they may play a role in glyphosate resistance. TMEM2-related proteins regulate levels of hyaluronic acid (HA) (Yamamoto et al., 2017). There is evidence suggesting HA's protective role against oxidative damage (Berdiaki et al., 2023). PKs are integral to stress tolerance in plants, with evidence suggesting their role in abiotic stress resistance (Majeed et al., 2023). Second, gene ontology (GO) terms associated with all variants in the final lists were compiled. Protein kinase activity (molecular function ontology) and membranes (cellular component ontology) were

consistently found in all selective histories. Thus, analysis of the reduced sets of genes, together with the GO terms list, suggested that glyphosate resistance may be partially mediated by pre-existing stress response pathways in *C. reinhardtii*, implying a potential role for non-target-site resistance (NTSR) mechanisms. Herbicide metabolism can lead to oxidative stress, highlighting the relationship between detoxification and stress response. NTSR mechanisms are part of a plant stress response (Délye and Christophe, 2013b). This finding contrasts with the assumption based on observed phenotype, that glyphosate resistance in *C. reinhardtii* was mediated by TSR mechanism (Lagator et al., 2012; Vogwill et al., 2012). Although *C. reinhardtii* may rely more on non-target site resistance (NTSR) mechanisms like enhanced efflux due to its unicellular nature, only limited genetic evidence for transporter-related resistance was observed in this study (a single replicate with a variant in *Cre09.g400950* annotated as encoding for an MFS protein). While stress-induced cell aggregation could contribute to glyphosate tolerance (de Carpentier et al., 2022), this phenotype was not detected experimentally, despite mutations in a Pherophorin homolog (*Cre14.g610700)* being present in all SR experiment replicates.

The impact of selective history on the genomic changes putatively linked to the GR phenotype was also investigated and seem to affect the distribution of genomic variants and the degree of genomic convergence across experimental replicates, suggesting convergent glyphosate resistant phenotype observed may be supported by divergent genomic changes. However, partial overlap in gene models impacted by variants was observed in the SR and the VR-quick selective histories, while greater heterogeneity was noted under intermediate selective history. This observation suggests that under the more stringent selective conditions, a small set of variants may play a critical role in glyphosate adaptation in *C. reinhardtii*.

6.2 Study limitations

Here I discuss the limitations to the work presented in this thesis. Despite *C. reinhardtii* being a good model for experimental evolution of herbicide resistance, whole genome analysis is not without its challenges: the genome is GC rich (64%), relatively large at 111Mb (Merchant et al., 2007) and contains numerous active transposable elements (Kim et al., 2006). However, since our results suggest the implication of NTSR glyphosate resistance mechanisms it remains a simple and interesting model to study this potentially complex, polygenic trait under tightly controlled conditions. The quality of the reference genome assembly and strain differences were a major limitation prior to the recent release of the highly contiguous reference genome assembled from reads from our laboratory strain by O'Donnell' et al (2020). Although CC-1690 and CC-503 strains are relatively closely related and most of the genome is identical by descent, they exhibit significant levels of polymorphism. The genome of CC-1690 features large duplications on chromosome 13, 12 new candidate genes and ~61,480 SNP when compared to the reference genome (Flowers et al., 2015). This led to practical difficulties in construction of an appropriate variant calling and variant filtering pipeline, imposing a time constraint on performing subsequent data analysis on the variants obtained. In hindsight, generating a high quality assembly of our lab strain at the outset would have helped avoid these challenges and would have improved variant calling by further reducing genome differences between the reference and the samples (Payne et al., 2023). In the SR selection experiment, clonality of different lines used to evolve resistance could not be assumed; this could explain some of the differences observed between replicates. This was addressed in the VR selection experiment by starting the selection from a single clone. This implies that both selection experiments

were not strictly comparable, limiting the inferences that can be drawn by comparing results for SR replicates or comparing results from SR samples and VR-quick. For example, there was little convergence observed between the GO terms of the SR and VR-quick samples. This could be due to either discrepancies in the selection experiment set up or the fact that too few generations underwent selection to observe high genomic convergence. The VR experiment could be replicated using a different culturing method to allow timing the rate of glyphosate dose increase to the number of generations under selection, allowing more meaningful comparisons between treatments. This might either be achieved by the use of a ratchet protocol (Melero-Jiménez et al., 2021) or continuous flow cultures (Hansson et al., 2022, 2024).

Investigation of the genomic changes putatively linked to glyphosate resistance was largely limited to the analysis of small variants. Further work on structural variants may provide a more complete understanding of the impact of glyphosate selection and selective history at the genomic level (Johnson et al., 2020), specifically on the presence of gene copy number variation (CNV). Although analysis of *EPSPS* copy number did not detect any likely role in resistance here, analysis of *EPSPS* differential expression may further inform on the implication of glyphosate TSR mechanisms and should be included in a future experiment.

Despite efforts to optimise the variant calling pipeline and downstream analysis of candidate variant calls through annotation approaches, there were still too many variants for further functional validation to be realistic. The final variant list might still contain false positives (FPs). My dataset contained a higher number of insertions than expected, based on the findings of Ness et al. (2015). These insertions could be false positives, possibly due to limitations in the current variant-calling pipeline. However, it is also plausible that they represent true variants, as stress conditions—such as salt stress—have been shown to increase the rate of indel mutations in *C. reinhardtii* (Hasan et al., 2022). This raises the possibility that the elevated insertion rate observed here may be a consequence of glyphosate-induced stress.To address this and further reduce potential FPs, future studies could refine the variant-calling pipeline by adjusting parameters for indel detection. Additionally, the presence of microsatellites in the final list of variants also raises concerns about FPs, as these repetitive DNA sequences are prone to high mutation rates due to DNA polymerase slippage (Kelkar et al., 2010). Despite these challenges, microsatellites may play a role in adaptation (Haasl & Payseur, 2013; K. Zhou et al., 2014)), making it inappropriate to exclude these sites without further investigation. GO term enrichment analysis might have helped reduce the list to fewer candidate variants but could not be performed due to lack of statistical power. Increasing sample size in any future studies may be beneficial in this regard.

The absence of results for the slow selective history limited interpretation on the impact of selective history on genomic changes. However good quality sequencing data is available and the analysis can be repeated to include this treatment.

Lastly, functional validation of variants was carried out for lack of an adequate final variant list as a consequence from the limitations discussed above. The next section will therefore discuss further work required to obtain this final variant list to take forward to functional validation.

6.3 New questions and further work

Further work to improve the current variant calling and variant filtering pipeline to obtain candidate variants for functional validation from this genome-wide approach is still currently limited by the lack of bioinformatic methods to achieve such goals. Future methods developments might require the combination of various methods for candidate variant detection. Multi-omics integration, by combining data from various omics layers (genomics, transcriptomic, proteomics and metabolomics) is an integrative approach that would enhance the accuracy of candidate gene identification (Abdullah-Zawawi et al., 2022). Once candidate variants are identified, functional validation is required to determine the causal relationship between the identified variants and their biological effects. In *C. reinhardtii,* methods such as CRISPR/Cas9 gene editing (Ghribi et al., 2020), RNA interference (Cerutti et al., 2011), and gene overexpression (Hema et al., 2007) can be used for functional validation of gene function.

Once the causal relationship between variants and glyphosate resistant phenotype in *C. reinhardtii* are established, examination of different genomic variants for their role in explaining the observed variance in glyphosate resistance associated fitness cost in the populations evolved under the intermediate and slow selective histories could be examined. Additionally, it would be possible to track the emergence and fixation or loss in populations over the course of a selection experiment, or even create various glyphosate resistant mutant lines in a homogeneous genetic background and subject them to competition in presence or absence of a selective pressure.

Further work aiming to link genomic changes to an observed glyphosate resistance phenotype in *C. reinhardtii* could also involve investigation over multiple time points during selection and track genomic changes over more generations in a similar way the long term evolution experiment started by Lenski in 1988 and running over 65,000 generations (Lenski, 2017b). This would be specifically helpful in investigating the effects of selective history on convergent or divergent evolution or replicates at the phenotypic and genomic levels.

The findings from this study provide important insights into the potential indirect impacts of glyphosate on microbial communities, particularly in freshwater systems. While glyphosate's effects on soil and aquatic microbiota are known to be highly context-dependent, with studies reporting both transient and significant shifts in community composition (Bueno de Mesquita et al., 2023; Newman et al., 2016), the evolutionary outcomes observed in *C. reinhardtii* suggest an additional layer of ecological complexity. Specifically, the evolution of resistance via non-target-site mechanisms implies that glyphosate exposure may select for a broad array of genetic responses, potentially altering microbial functional diversity. This aligns with ecotoxicological evidence that herbicide-resistant *C. reinhardtii* populations often exhibit fitness trade-offs, such as reduced growth and photosynthetic performance (Melero-Jimenez et al., 2021), which could diminish primary production. Furthermore, given that glyphosate exposure has been linked to shifts in algal species composition and ecosystem dynamics (Saxton et al., 2011; Zhang et al., 2023), the observed variability and genomic divergence in resistance evolution across replicates reinforce concerns about long-term impacts on microbial ecosystems. Taken together, these results underscore the importance of considering evolutionary trajectories and resistance mechanisms when assessing glyphosate’s indirect ecological effects in the rhizosphere and aquatic environments which could be the focus of future studies.

6.4 Conclusions

This study demonstrates that glyphosate resistance likely evolves through diverse mechanisms in *C. reinhardtii* under different selection regimes, influencing both phenotypic and genomic changes.

Findings highlight both the repeatability and variability of evolution in the context of glyphosate resistance. Consistent evolution of resistance to glyphosate doses up to 1 MIC was observed. Evolution of glyphosate resistant phenotypes was generally repeatable, aligning with prior studies. However, selective history shaped the resistance associated fitness costs and evolutionary dynamics. Quick dose escalation imposed strong selective pressure, reducing population size and delaying resistance evolution, while gradual dose increases led to more stable populations but higher fitness costs. There is also evidence that selective history impacts genomic changes and the degree of genomic convergence between replicates. Under quick selective pressure, a smaller set of variants appeared critical, suggesting convergent evolution at the genomic level. In contrast, intermediate selective histories showed greater genomic heterogeneity, implying divergent evolutionary paths to achieve similar phenotypic outcomes.

Genome-variant analysis revealed that *EPSPS* mutations or gene copy number variation (TSR mechanisms) were not present, suggesting variants could support other known mechanism such as *EPSPS* over expression (TSR mechanism) or non-target-site resistance (NTSR) mechanisms. There was preliminary evidence for the involvement of variants associated with stress response pathways, including protein kinases and TMEM2-related proteins. These findings suggest that glyphosate resistance might leverage pre-existing cellular stress tolerance pathways, a novel finding in this context. This challenges the assumption that TSR mechanisms dominate herbicide resistance in this species.

Innovative aspects of the work presented here lie in its combination of experimental evolution with whole-genome sequencing (E&R approach) to investigate glyphosate resistance of *C. reinhardtii* populations. The research developed custom pipelines for variant calling and filtering tailored to the unique challenges of the *C. reinhardtii* genome, an important first step towards the identification of candidate variants linked to glyphosate resistance.

While the research advances understanding of glyphosate resistance evolution, there is scope for further optimisation of the approaches developed here. In addition, the existing data sets will be a useful resource for future analyses beyond the scope of the current study, such as a more comprehensive investigation of structural variants. Future work should integrate multi-omics approaches, validate candidate variants, and examine genomic changes across longer evolutionary timelines to better elucidate the genetic basis of resistance.

References

Abdullah-Zawawi, M.R., Govender, N., Harun, S., Muhammad, N.A.N., Zainal, Z., and Mohamed-Hussein, Z.A. (2022). Multi-Omics Approaches and Resources for Systems-Level Gene Function Prediction in the Plant Kingdom. Plants 11.

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., and 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, 1–12.

Andersson, D.I., and Hughes, D. (2010). Antibiotic resistance and its cost: Is it possible to reverse resistance? Nat Rev Microbiol 8, 260–271.

Andrews, S. (2010). FastQC A Quality Control tool for High Throughput Sequence Data. Retreived from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Antier, C., Kudsk, P., Reboud, X., Ulber, L., Baret, P. V., and Messéan, A. (2020). Glyphosate use in the European agricultural sector and a framework for its further monitoring. Sustainability (Switzerland) 12, 1–22.

Atin, S., Rajnikant, S., Tapas, B., Timsy, B., and Ranjana, P. (2017). Fosfomycin resistance in Acinetobacter baumannii is mediated by efflux through a major facilitator superfamily (MFS) transporter-AbaF. J. Antimicrob. Chemother. *72*, 68–74.

Baek, Y., Bobadilla, L.K., Giacomini, D.A., Montgomery, J.S., Murphy, Brent.P., and Tranel, P.J. (2021). Evolution of Glyphosate-resistants weeds. In Reviews of Environmental Contamination and Toxicology, Dr. James B. Knaak, ed. (Srpinger, Cham), pp. 93–128.

Baerson, S.R., Rodriguez, D.J., Tran, M., Feng, Y., Biest, N.A., and Dill, G.M. (2002). Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. Plant Physiol 129, 1265–1275.

Barbitoff, Y.A., Matveenko, A.G., Matiiv, A.B., Maksiutenko, E.M., Moskalenko, S.E., Drozdova, P.B., Polev, D.E., Beliavskaia, A.Y., Danilov, L.G., Predeus, A. V., et al. (2021). Chromosome-level genome assembly and structural variant analysis of two laboratory yeast strains from the Peterhof Genetic Collection lineage. G3: Genes, Genomes, Genetics 11.

Barrick, J.E., and Lenski, R.E. (2013). Genome dynamics during experimental evolution. Nat Rev Genet 14, 827–839.

Bass, C., and Field, L.M. (2011). Gene amplification and insecticide resistance. Pest Manag Sci 67, 886–890.

Battaglin, W.A., Meyer, M.T., Kuivila, K.M., and Dietze, J.E. (2014). Glyphosate and its degradation product AMPA occur frequently and widely in U.S. soils, surface water, groundwater, and precipitation. J Am Water Resour Assoc 50, 275–290.

Baylis, A.D. (2000). Why glyphosate is a global herbicide: Strengths, weaknesses and prospects. Pest Manag Sci 56, 299–308.

Beckie, H.J. (2011). Herbicide-resistant weed management: Focus on glyphosate. Pest Manag Sci 67, 1037–1048.

Behringer, M.G., Choi, B.I., Miller, S.F., Doak, T.G., Karty, J.A., Guo, W., and Lynch, M. (2018). Escherichia coli cultures maintain stable subpopulation structure during long-term evolution. Proc Natl Acad Sci U S A 115, E4642–E4650.

Bell, G. (1990a). The ecology and genetics of fitness in Chlamydomonas. I. Genotype-by-environment interaction among pure strains. Biological Sciences Proc. R. Soc. Lond. B 240, 295–321.

Bell, G. (1990b). The Ecology and Genetics of Fitness in Chlamydomonas. II. The Properties of Mixtures of Strains. Proceedings of the Royal Society B: Biological Sciences 240, 323–350.

Bell, G. (1997). Experimental evolution in Chlamydomonas. I. Short-term selection in uniform and diverse environments. Heredity (Edinb) 78, 490–497.

Bell, G., and Collins, S. (2008). Adaptation, extinction and global change. Evol Appl 1, 3–16.

Bell, G., and Gonzalez, A. (2011). Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. Science (1979) 332, 1327–1330.

Bell, G., and Reboud, X. (1997). Experimental evolution in Chlamydomonas II. Genetic variation in strongly contrasted environments. Heredity (Edinb) 78, 1–11.

Benbrook, C.M. (2016). Trends in glyphosate herbicide use in the United States and globally. Environ Sci Eur 28, 1–15.

Bentley, R. (1990). The sikimate pathway- a metabolic tree with many branches. Biochemistry and Molecular Biology 25, 307–383.

Berdiaki, A., Neagu, M., Spyridaki, I., Kuskov, A., Perez, S., and Nikitovic, D. (2023). Hyaluronan and Reactive Oxygen Species Signaling—Novel Cues from the Matrix? Antioxidants 12, 1–26.

Blaby, I.K., Blaby-Haas, C.E., Tourasse, N., Hom, E.F., Lopez, D., Aksoy, M., Grossman, A., Umen, J., Dutcher, S., Porter, M., et al. (2014). The Chlamydomonas genome project: a decade on. Trends Plant Sci 19, 672–680.

Blanquart, F. (2019). Evolutionary epidemiology models to predict the dynamics of antibiotic resistance. Evol. Appl. *12*, 365–383.

Blum, M., Andreeva, A., Florentino, L.C., Chuguransky, S.R., Grego, T., Hobbs, E., Pinto, B.L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., et al. (2025). InterPro: the protein sequence classification resource in 2025. Nucleic Acids Res 53, D444–D456.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Böndel, K.B., Kraemer, S.A., Samuels, T.S., McClean, D., Lachapelle, J., Ness, R.W., Colegrave, N., and Keightley, P.D. (2019). Direct inference of the distribution of fitness effects of spontaneous mutations in Chlamydomonas reinhardtii. BioRxiv 571018.

Boocock, M.R., and Coggins, J.R. (1983). Kinetics of 5-enolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate. FEBS Lett 154, 127–133.

Boulé, J.B., and Zakian, V.A. (2006). Roles of Pif1-like helicases in the maintenance of genomic stability. Nucleic Acids Res 34, 4147–4153.

Bradshaw, L.D., Padgette, S.R., Kimball, S.L., and Wells, B.H. (1997). Mini-Review / Commentary Perspectives on Glyphosate Resistance ’. Weed Science Society of America 11, 189–198.

Broad Institute (2019). Picard Toolkit. Retreived from https://broadinstitute.github.io/picard/.

Brockhurst, M.A., Colegrave, N., and Rozen, D.E. (2011). Next-generation sequencing as a tool to study microbial evolution. Mol Ecol 20, 972–980.

Bruggeman, A.J., Kuehler, D., and Weeks, D.P. (2014). Evaluation of three herbicide resistance genes for use in genetic transformations and for potential crop protection in algae production. Plant Biotechnol J 12, 894–902.

Van Bruggen, A.H.C., He, M.M., Shin, K., Mai, V., Jeong, K.C., Finckh, M.R., and Morris, J.G. (2018). Environmental and health effects of the herbicide glyphosate. Science of the Total Environment 616–617, 255–268.

Buckling, A., Brockhurst, M.A., Travisano, M., and Rainey, P.B. (2007). Experimental adaptation to high and low quality environments under different scales of temporal variation. J Evol Biol 20, 296–300.

Buckling, A., Craig Maclean, R., Brockhurst, M.A., and Colegrave, N. (2009). The Beagle in a bottle. Nature 457, 824–829.

Bueno de Mesquita, C.P., Solon, A.J., Barfield, A., Mastrangelo, C.F., Tubman, A.J., Vincent, K., Porazinska, D.L., Hufft, R.A., Shackelford, N., Suding, K.N., et al. (2023). Adverse impacts of Roundup on soil bacteria, soil chemistry and mycorrhizal fungi during restoration of a Colorado grassland. Appl. Soil Ecol. *185*, 104778.

Burger, R., and Lynch, M. (1995). Evolution and extinction in a changing environment: A quantitative-genetic analysis. Evolution (N Y) 49, 151–163.

Burke, M.K., Liti, G., and Long, A.D. (2014). Standing genetic variation drives repeatable experimental evolution in outcrossing populations of saccharomyces cerevisiae. Mol Biol Evol 31, 3228–3239.

Bvindi, C., Lee, S., Tang, L., Mickelbart, M. V., Li, Y., and Mengiste, T. (2022). Improved pathogen and stress tolerance in tomato mutants of SET domain histone 3 lysine methyltransferases. New Phytologist 235, 1957–1976.

Caetano-Anolles, D. (2024). GATK Best Practices for Structural Variation Discovery on Single Samples. Retreived from https://gatk.broadinstitute.org/hc/en-us/articles/9022653744283-GATK-Best-Practices-for-Structural-Variation-Discovery-on-Single-Samples.

Campbell, S.A., Richards, T.A., Mui, E.J., Samuel, B.U., Coggins, J.R., McLeod, R., and Roberts, C.W. (2004). A complete shikimate pathway in Toxoplasma gondii: an ancient eukaryotic innovation. Int J Parasitol 34, 5–13.

de Carpentier, F., Maes, A., Marchand, C.H., Chung, C., Durand, C., Crozet, P., Lemaire, S.D., and Danon, A. (2022). How abiotic stress-induced socialization leads to the formation of massive aggregates in Chlamydomonas. Plant Physiol 190, 1927–1940.

Carroll, S.P., Hendry, A.P., Reznick, D.N., and Fox, C.W. (2007). Evolution on ecological time-scales. Funct Ecol 21, 387–393.

De Carvalho, L.B., Alves, P.L.D.C.A., González-Torralva, F., Cruz-Hipolito, H.E., Rojano-Delgado, A.M., De Prado, R., Gil-Humanes, J., Barro, F., and Luque De Castro, M.D. (2012). Pool of resistance mechanisms to glyphosate in digitaria insularis. J Agric Food Chem 60, 615–622.

Castle, L.A., Siehl, D.L., Gorton, R., Patten, P.A., Chen, Y.H., Bertain, S., Cho, H.J., Duck, N., Wong, J., Liu, D., et al. (2004). Discovery and directed evolution of a glyphosate tolerance gene. Science 304, 1151–1154.

Cerutti, H., Ma, X., Msanne, J., and Repas, T. (2011). RNA-mediated silencing in Algae: Biological roles and tools for analysis of gene function. Eukaryot Cell 10, 1164–1172.

Chen, C.M., Ye, Q.Z., Zhu, Z., Wanner, B.L., and Walsh, C.T. (1990). Molecular biology of carbon-phosphorus bond cleavage: Cloning and sequencing of the phn (psiD) genes involved in alkylphosphonate uptake and C-P lyase activity in Escherichia coli B. Journal of Biological Chemistry 265, 4461–4471.

Chevereau, G., Dravecká, M., Batur, T., Guvenek, A., Ayhan, D.H., Toprak, E., and Bollenbach, T. (2015). Quantifying the Determinants of Evolutionary Dynamics Leading to Drug Resistance. PLoS Biol 13, 1–18.

Chevin, L.M., Lande, R., and Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. PLoS Biol 8.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Ruden, D.M., and Lu, X. (2012a). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 6, 80–92.

Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M., and Lu, X. (2012b). Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 3, 35.

Cocker, K.M., Moss, S.R., and Coleman, J.O.D. (1999). Multiple mechanisms of resistance to fenoxaprop-P-ethyl in United Kingdom and other European populations of herbicide-resistant Alopecurus myosuroides (black-grass). Pestic Biochem Physiol 65, 169–180.

Collavo, A., and Sattin, M. (2012). Resistance to glyphosate in Lolium rigidum selected in Italian perennial crops: Bioevaluation, management and molecular bases of target-site resistance. Weed Res 52, 16–24.

Collins, S. (2011). Many Possible Worlds: Expanding the Ecological Scenarios in Experimental Evolution. Evol Biol 38, 3–14.

Collins, S., and De Meaux, J. (2009). Adaptation to different rates of environmental change in chlamydomonas. Evolution (N Y) 63, 2953–2965.

Collins, S., De Meaux, J., and Acquisti, C. (2007). Adaptive walks toward a moving optimum. Genetics 176, 1089–1099.

Comont, D., Hicks, H., Crook, L., Hull, R., Cocciantelli, E., Hadfield, J., Childs, D., Freckleton, R., and Neve, P. (2019a). Evolutionary epidemiology predicts the emergence of glyphosate resistance in a major agricultural weed. New Phytologist 223, 1584–1594.

Comont, D., Knight, C., Crook, L., Hull, R., Beffa, R., and Neve, P. (2019b). Alterations in life-history associated with non-target-site herbicide resistance in Alopecurus myosuroides. Front Plant Sci 10, 1–16.

Condon, C., Cooper, B.S., Yeaman, S., and Angilletta, M.J.J. (2014). Temporal variation favors the evolution of generalists in experimental populations of Drosophila melanogaster. Evolution (N Y) 68, 720–728.

Cooper, T.F., Rozen, D.E., and Lenski, R.E. (2003). Parallel changes in gene expression after 20,000 generations of evolution in Escherichia coli. Proc Natl Acad Sci U S A 100, 1072–1077.

Costelloe, C., Metcalfe, C., Lovering, A., Mant, D., and Hay, A.D. (2010). Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: Systematic review and meta-analysis. BMJ *340*, 1120.

Damalas, C.A., and Koutroubas, S.D. (2024). Herbicide resistance evolution, fitness cost, and the fear of the superweeds. Plant Science 339, 111934.

Day, T., and Read, A.F. (2016). Does High-Dose Antimicrobial Chemotherapy Prevent the Evolution of Resistance? PLoS Comput. Biol. *12*, 1–20.

Decottignies, A., Rogers, B., Kolaczkowski, M., Carvajal, E., Balzi, E., Conseil, G., Niimi, K., Di Pietro, A., Monk, B.C., and Goffeau, A. (2002). The pleiotropic drug ABC transporters from Saccharomyces cerevisiae.

Délye, C., and Christophe (2013). Unravelling the genetic bases of non-target-site-based resistance (NTSR) to herbicides: A major challenge for weed science in the forthcoming decade. Pest Manag Sci 69, 176–187.

Délye, C., Zhang, X.Q., Michel, S., Matéjicek, A., and Powles, S.B. (2005). Molecular bases for sensitivity to acetyl-coenzyme a carboxylase inhibitors in black-grass. Plant Physiol 137, 794–806.

Délye, C., Jasieniuk, M., and Le Corre, V. (2013). Deciphering the evolution of herbicide resistance in weeds. Trends in Genetics 29, 649–658.

Denamur, E., Tenaillon, O., Deschamps, C., Skurnik, D., Ronco, E., Gaillard, J.L., Picard, B., Branger, C., and Matic, I. (2005). Intermediate mutation frequencies favor evolution of multidrug resistance in Escherichia coli. Genetics 171, 825–827.

Deng, W., Duan, Z., Li, Y., Peng, C., and Yuan, S. (2022). Multiple Resistance Mechanisms Involved in Glyphosate Resistance in Eleusine indica. Plants 11.

Deng, Z.L., Dhingra, A., Fritz, A., Götting, J., Münch, P.C., Steinbrück, L., Schulz, T.F., Ganzenmüller, T., and McHardy, A.C. (2021). Evaluating assembly and variant calling software for strain-resolved analysis of large DNA viruses. Brief Bioinform 22, 1–12.

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., Chan, K.M.A., et al. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change. Science (1979) 1327.

Dubey, K.K., Kumar, P., Singh, P.K., and Shukla, P. (2014). Exploring Prospects of Monooxygenase-Based Biocatalysts in Xenobiotics. In Microbial Biodegradation and Bioremediation, (Elsevier Inc.), pp. 580–616.

Duke, S.O., and Powles, S.B. (2008). Glyphosate: a once-in-a-century herbicide. Pest Manag Sci 63, 1100–1106.

Duke, S.O., Heap, I., Tranel, P.J., and Bobadilla, L.K. (2021). How many ways can nature kill the goose that laid the golden egg? – the many mechanisms of evolved glyphosate resistance. Outlooks on Pest Management 32, 197–202.

Dupuis, S., and Merchant, S.S. (2023). Chlamydomonas reinhardtii: a model for photosynthesis and so much more. Nat Methods 20, 1441–1442.

Edwards, C.B., Jordan, D.L., DKOwen, M., MDixon, P., Young, B.G., GWilson, R., CWellerf, S., and Shawa, D.R. (2014). Benchmark study on glyphosate-resistant cropsystems in the United States. Economicsof herbicide resistance management practicesin a 5 year field-scale study. Pest Manag Sci 70, 1924–1929.

Elena, S.F., and Lenski, R.E. (2003). Microbial genetics: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet 4, 457–469.

Elena, S.F., Wilke, C.O., Ofria, C., and Lenski, R.E. (2007). Effects of population size and mutation rate on the evolution of mutational robustness. Evolution (N Y) 61, 666–674.

Erickson, J.M., Rahire, M., Bennoun, P., Delepelaire, P., Diner, B., and Rochaix, J.-D. (1984). Herbicide resistance in Chlamydomonas reinhardtii results from a mutation in the chloroplast gene for the 32-kilodalton protein of photosystem II. Proceedings of the National Academy of Sciences 81, 3617–3621.

Erickson, J.M., Pfister, K., Rahire, M., Togasaki, R.K., and Mets, L. (1989). Molecular and Biophysical Analysis of Herbicide- Resista n t M u t a n t s of Chlam ydomonas reínhardtií : Structure-Function Relationship of the Photosystem II D 1 Poly peptide. 1, 361–371.

Van Etten, M., Lee, K.M., Chang, S.M., and Baucom, R.S. (2020). Parallel and nonparallel genomic responses contribute to herbicide resistance in Ipomoea purpurea, a common agricultural weed.

Fedtke, C. (1991). Mode of action studies with mefenacet. Pestic Sci 33, 421–426.

Feng, D., Soric, A., and Boutin, O. (2020). Treatment technologies and degradation pathways of glyphosate: A critical review. Science of the Total Environment 742, 140559.

Ferriere, R., and Legendre, S. (2013). Eco-evolutionary feedbacks, adaptive dynamics and evolutionary rescue theory. Philosophical Transactions of the Royal Society B: Biological Sciences 368.

Firdous, S., Iqbal, S., and Anwar, S. (2020). Optimization and modeling of glyphosate biodegradation by a novel Comamonas odontotermitis P2 through response surface methodology. Pedosphere 30, 618–627.

Fitzgibbon, J., and Braymer, H.D. (1988). Phosphate starvation induces uptake of glyphosate by Pseudomonas sp. strain PG2982. Appl Environ Microbiol 54, 1886–1888.

Fitzgibbon, J.E., and Braymer, H.D. (1990). Cloning of a gene from Pseudomonas sp. strain PG2982 conferring increased glyphosate resistance. Appl Environ Microbiol 56, 3382–3388.

Flowers, J.M., Hazzouri, K.M., Pham, G.M., Rosas, U., Bahmani, T., Khraiwesh, B., Nelson, D.R., Jijakli, K., Abdrabu, R., Harris, E.H., et al. (2015). Whole-genome resequencing reveals extensive natural variation in the model green alga chlamydomonas reinhardtii. Plant Cell 27, 2353–2369.

Fu, S., Huang, J., Chen, Z., and Xia, J. (2019). C2 domain plays critical roles in localization of novel C2 domain-containing protein OsC2DP. Plant Signal Behav 14.

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., et al. (2010). Gene amplification confers glyphosate resistance in Amaranthus palmeri. Proc Natl Acad Sci U S A 107, 1029–1034.

Gaines, T.A., Patterson, E.L., and Neve, P. (2019). Molecular mechanisms of adaptive evolution revealed by global selection for glyphosate resistance. New Phytologist 223, 1770–1775.

Gaines, T.A., Duke, S.O., Morran, S., Rigon, C.A.G., Tranel, P.J., Küpper, A., and Dayan, F.E. (2020). Mechanisms of evolved herbicide resistance. J Biol Chem 295, 10307–10330.

Galeano, E., Barroso, A.A.M., Vasconcelos, T.S., López-Rubio, A., Albrecht, A.J.P., Victoria Filho, R., and Carrer, H. (2016). EPSPS variability, gene expression, and enzymatic activity in glyphosate-resistant biotypes of Digitaria insularis. Genetics and Molecular Research 15, 1–15.

Gallaher, S.D., Fitz-Gibbon, S.T., Glaesener, A.G., Pellegrini, M., and Merchanta, S.S. (2015). Chlamydomonas genome resource for laboratory strains reveals a mosaic of sequence variation, identifies true strain histories, and enables strain-specific studies. Plant Cell 27, 2335–2352.

Gallaher, S.D., Fitz-Gibbon, S.T., Strenkert, D., Purvine, S.O., Pellegrini, M., and Merchant, S.S. (2018). High-throughput sequencing of the chloroplast and mitochondrion of Chlamydomonas reinhardtii to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. Plant Journal 93, 545–565.

Galloway, R.E., and Mets, L.J. (1984). Atrazine, Bromacil, and Diuron Resistance in Chlamydomonas. Plant Physiol 74, 469–474.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F., and Conesa, A. (2012). Qualimap: Evaluating next-generation sequencing alignment data. Bioinformatics 28, 2678–2679.

Garland, T., and Rose, M.R. (2009). Experimental evolution: Concepts, methods, and applications of selection experiments.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. ArXiv preprint, 1–9.

Ge, X., D'Avignon, D.A., Ackerman, J.J., Duncan, B., Spaur, M.B., and Sammons, R.D. (2011). Glyphosate-resistant horseweed made sensitive to glyphosate: Low-temperature suppression of glyphosate vacuolar sequestration revealed by 31P NMR. Pest Manag Sci 67, 1215–1221.

Gerakari, M., Cheimona, N., Tani, E., Travlos, I., Chachalis, D., Loddo, D., Mathiassen, S.K., Gitsopoulos, T.K., Scarabel, L., Panozzo, S., et al. (2022). Biochemical and rapid molecular analyses to identify glyphosate resistance in lolium spp. Agronomy 12, 1–12.

Ghribi, M., Nouemssi, S.B., Meddeb-Mouelhi, F., and Desgagné-Penix, I. (2020). Genome editing by CRISPR-Cas: A game change in the genetic manipulation of chlamydomonas. Life 10, 1–21.

Giacomini, D., Westra, P., and Ward, S.M. (2014). Impact of Genetic Background in Fitness Cost Studies: An Example from Glyphosate-Resistant Palmer Amaranth. Weed Sci 62, 29–37.

Giacomini, D.A., Patterson, E.L., Küpper, A., Beffa, R., Gaines, T.A., and Tranel, P.J. (2020). Coexpression clusters and allele-specific expression in metabolism-based herbicide resistance. Genome Biol Evol 12, 2267–2278.

Giesy, J.P., Dobson, S., and Solomon, K.R. (2000). Ecotoxicological risk assessment for Roundup® herbicide. Rev Environ Contam Toxicol 167, 35–120.

Gill, S.S., Anjum, N.A., Hasanuzzaman, M., Gill, R., Trivedi, D.K., Ahmad, I., Pereira, E., and Tuteja, N. (2013). Glutathione and glutathione reductase: A boon in disguise for plant abiotic stress defense operations. Plant Physiology and Biochemistry 70, 204–212.

Gonçalves, B.B., Giaquinto, P.C., Silva, D. dos S., Neto, C. de M. e S., Lima, A.A. de, Darosci, A.A.B., Portinho, J.L., Carvalho, W.F., and Rocha, T.L. (2020). Ecotoxicology of Glyphosate-Based Herbicides on Aquatic Environment. In Organic Pollutants, (IntechOpen), p.

Gonzalez, A., and Bell, G. (2013). Evolutionary rescue and adaptation to abrupt environmental change depends upon the history of stress. Philosophical Transactions of the Royal Society B: Biological Sciences 368.

Gonzalez, A., Ronce, O., Ferriere, R., and Hochberg, M.E. (2013). Evolutionary rescue: An emerging focus at the intersection between ecology and evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20120404.

González-Torralva, F., Gil-Humanes, J., Barro, F., Brants, I., and De Prado, R. (2012). Target site mutation and reduced translocation are present in a glyphosate-resistant Lolium multiflorum Lam. biotype from Spain. Plant Physiology and Biochemistry 58, 16–22.

Good, B.H., McDonald, M.J., Barrick, J.E., Lenski, R.E., and Desai, M.M. (2017). The dynamics of molecular evolution over 60,000 generations. Nature 551, 45–50.

Gorter, F.A., Aarts, M.M.G.G., Zwaan, B.J., de Visser, J.A.G.M., Visser, J.A.G.M. De, and de Visser, J.A.G.M. (2016). Dynamics of adaptation in experimental yeast populations exposed to gradual and abrupt change in heavy metal concentration. American Naturalist 187, 110–119.

Gorter, F.A., Derks, M.F.L.L., Van Den Heuvel, J., Aarts, M.G.M.M., Zwaan, B.J., De Ridder, D., De Visser, J.A.G.M., Heuvel, J. Van Den, Aarts, M.G.M.M., Zwaan, B.J., et al. (2017). Genomics of Adaptation Depends on the Rate of Environmental Change in Experimental Yeast Populations. Mol Biol Evol 34, 2613–2626.

Graham, L.D., Gillies, F.M., and Coggins, J.R. (1993). Over-expression of the yeast multifunctional arom protein. BBA - Gene Structure and Expression 1216, 417–424.

Gray, J.C., and Goddard, M.R. (2012). Sex enhances adaptation by unlinking beneficial from detrimental mutations in experimental yeast populations. BMC Evol Biol 12.

Gressel, J. (2009). Evolving understanding of the evolution of herbicide resistance. Pest Manag Sci 65, 1164–1173.

Gressel, J. (2011). Low pesticide rates may hasten the evolution of resistance by increasing mutation frequencies. Pest Manag Sci 67, 253–257.

Guo, J., Song, X., Zheng, C., Sun, S., Zhuang, B., and Tao, B. (2021). Transcriptome analysis and identification of candidate genes involved in glyphosate resistance in the fungus Fusarium verticillioides. J Environ Sci Health B 56, 658–669.

Haasl, R.J., and Payseur, B.A. (2013). Microsatellites as targets of natural selection. Mol Biol Evol 30, 285–298.

Hairston, N.G., Ellner, S.P., Geber, M.A., Yoshida, T., and Fox, J.A. (2005). Rapid evolution and the convergence of ecological and evolutionary time. Ecol Lett 8, 1114–1127.

Hallmann, A. (2006). The pherophorins: Common, versatile building blocks in the evolution of extracellular matrix architecture in Volvocales. Plant Journal 45, 292–307.

Han, H., Yu, Q., Widderick, M.J., and Powles, S.B. (2016). Target-site EPSPS Pro-106 mutations: Sufficient to endow glyphosate resistance in polyploid Echinochloa colona? Pest Manag Sci 72, 264–271.

Han, H., Vila-Aiub, M., Jalaludin, A., Yu, Q., and Powles, S.B. (2017). A double EPSPS gene mutation endowing glyphosate resistance shows a remarkably high resistance cost. Plant Cell Environ 40.

Hansson, E.M., Childs, D.Z., and Beckerman, A.P. (2022). Mesostats-A multiplexed, low-cost, do-ityourself continuous culturing system for experimental evolution of mesocosms. PLoS One 17, 1–15.

Hansson, E.M., Childs, D.Z., and Beckerman, A.P. (2024). Glyphosate resistance evolution to lethal and sublethal doses in chemostat populations of model organism Chlamydomonas reinhardtii. 1–20.

Harre, N.T., Nie, H., Jiang, Y., and Young, B.G. (2018). Differential antioxidant enzyme activity in rapid-response glyphosate-resistant Ambrosia trifida. Pest Manag Sci 74, 2125–2132.

Harris, E.H. (2008a). Chapter 8: Chlamydomonas in the laboratory. In The Chlamydomonas Source Book. Introduction to Chlamydomonas and Its Laboratory Use, (Elsevier), pp. 241–302.

Harris, E.H. (2008b). Origins of the major laboratory strains of C.reinhardtii. In The Chlamydomonas Source Book.Introduction to Chlamydomonas and Its Laboratory Use., (Elsevier), pp. 11–16.

Hartnett, M.E., Newcomb, J.R., and Hodson, R.C. (1987). Mutations in Chlamydomonas reinhardtii Conferring Resistance to the Herbicide Sulfometuron Methyl. Plant Physiol 85, 898–901.

Hasan, A.R., Lachapelle, J., El-Shawa, S.A., Potjewyd, R., Ford, S.A., and Ness, R.W. (2022). Salt stress alters the spectrum of de novo mutation available to selection during experimental adaptation of Chlamydomonas reinhardtii. Evolution (N Y) 76, 2450–2463.

Haydon, M.J., and Cobbett, C.S. (2007). A novel major facilitator superfamily protein at the tonoplast influences zinc tolerance and accumulation in Arabidopsis. Plant Physiol. *143*, 1705–1719.

Healy-Fried, M.L., Funke, T., Priestman, M.A., Han, H., and 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.

Heap, I. (2021). www.weedscience.org.

Heap, I. (2024). The International Herbicide-Resistant Weed Database.

Hema, R., Senthil-Kumar, M., Shivakumar, S., Chandrasekhara Reddy, P., and Udayakumar, M. (2007). Chlamydomonas reinhardtii, a model system for functional validation of abiotic stress responsive genes. Planta 226, 655–670.

Herron, M.D., and Doebeli, M. (2013). Parallel Evolutionary Dynamics of Adaptive Diversification in Escherichia coli. PLoS Biol 11.

Hertel, R., Gibhardt, J., Martienssen, M., Kuhn, R., and Commichau, F.M. (2021). Molecular mechanisms underlying glyphosate resistance in bacteria. Environ Microbiol 23, 2891–2905.

von der Heyde, B., and Hallmann, A. (2023). Cell Type-Specific Pherophorins of Volvox carteri Reveal Interplay of Both Cell Types in ECM Biosynthesis. Cells 12.

Hotta, Y., and Bassel, A. (1965). Molecular Size and Circularity of Dna in Cells of Mammals and Higher plants. Proceedings of the National Academy of Sciences of the United States Of 53, 356–362.

Hove-Jensen, B., Rosenkrantz, T.J., Zechel, D.L., and Willemoës, M. (2010). Accumulation of intermediates of the carbon-phosphorus lyase pathway for phosphonate degradation in phn mutants of Escherichia coli. J Bacteriol 192, 370–374.

Hove-Jensen, B., McSorley, F.R., and Zechel, D.L. (2011). Physiological role of phnP-specified phosphoribosyl cyclic phosphodiesterase in catabolism of organophosphonic acids by the carbon-phosphorus lyase pathway. J Am Chem Soc 133, 3617–3624.

Hove-Jensen, B., Zechel, D.L., and Jochimsen, B. (2014). Utilization of Glyphosate as Phosphate Source: Biochemistry and Genetics of Bacterial Carbon-Phosphorus Lyase. Microbiology and Molecular Biology Reviews 78, 176–197.

Huntscha, S., Stravs, M.A., Bühlmann, A., Ahrens, C.H., Frey, J.E., Pomati, F., Hollender, J., Buerge, I.J., Balmer, M.E., and Poiger, T. (2018). Seasonal Dynamics of Glyphosate and AMPA in Lake Greifensee: Rapid Microbial Degradation in the Epilimnion during Summer. Environ Sci Technol 52, 4641–4649.

IARC (2015). IARC Monographs Volume 112: evaluation of five organophosphate insecticides and herbicides (Lyon).

Jain, M., Aggarwal, S., Nagar, P., Tiwari, R., and Mustafiz, A. (2020). A D-lactate dehydrogenase from rice is involved in conferring tolerance to multiple abiotic stresses by maintaining cellular homeostasis. Sci Rep 10, 1–17.

James, S.W., Silflow, C.D., Stroom, P., and Lefebvre, P.A. (1993). A mutation in the α1-tubulin gene of Chlamydomonas reinhardtii confers resistance to anti-microtubule herbicides. J Cell Sci 106, 209–218.

Jander, G., Baerson, S.R., Hudak, J.A., Gonzalez, K.A., Gruys, K.J., and Last, R.L. (2003). Ethylmethanesulfonate saturation mutagenesis in Arabidopsis to determine frequency of herbicide resistance. Plant Physiol 131, 139–146.

Jansen, G., Barbosa, C., and Schulenburg, H. (2013). Experimental evolution as an efficient tool to dissect adaptive paths to antibiotic resistance. Drug Resistance Updates 16, 96–107.

Jasieniuk, M., Ahmad, R., Sherwood, A.M., Firestone, J.L., Perez-Jones, A., Lanini, W.T., Mallory-Smith, C., and Stednick, Z. (2008). Glyphosate-Resistant Italian Ryegrass (Lolium multiflorum) in California: Distribution, Response to Glyphosate, and Molecular Evidence for an Altered Target Enzyme . Weed Sci 56, 496–502.

Jerison, E.R., Kryazhimskiy, S., Mitchell, J.K., Bloom, J.S., Kruglyak, L., and Desai, M.M. (2017). Genetic variation in adaptability and pleiotropy in budding yeast. Elife 6.

Jessup, C.M., Kassen, R., Forde, S.E., Kerr, B., Buckling, A., Rainey, P.B., and Bohannan, B.J.M. (2004). Big questions, small worlds: Microbial model systems in ecology. Trends Ecol Evol 19, 189–197.

Johnson, N.A., Lemas, J., Montgomery, J., Gaines, T., and Patterson, E. (2020). Genomic structural variation and herbicide resistance.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: Genome-scale protein function classification. Bioinformatics 30, 1236–1240.

Jugulam, M., Niehues, K., Godar, A.S., Koo, D.H., Danilova, T., Friebe, B., Sehgal, S., Varanasi, V.K., Wiersma, A., Westra, P., et al. (2014). Tandem amplification of a chromosomal segment harboring 5-Enolpyruvylshikimate-3-Phosphate synthase locus confers glyphosate resistance in Kochia scoparia. Plant Physiol 166, 1200–1207.

Kanwischer, M., Porfirova, S., Bergmüller, E., and Dörmann, P. (2005). Alterations in tocopherol cyclase activity in transgenic and mutant plants of Arabidopsis affect tocopherol content, tocopherol composition, and oxidative stress. Plant Physiol 137, 713–723.

Kao, K.C., and Sherlock, G. (2008). Molecular characterization of clonal interference during adaptive evolution in asexual populations of Saccharomyces cerevisiae. Nat Genet 40, 1499–1504.

Kassen, R., and Bataillon, T. (2006). Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. 38, 484–488.

Kassen, R., and Bell, G. (1998). Experimental evolution in Chlamydomonas. IV. Selection in environments that vary through time at different scales. Heredity (Edinb) 80, 732–741.

Kawecki, Tadeusz.J., and Holt, Robert.D. (2002). Evolutionary consequences of asymmetric dispersal rates. Am Nat 160, 334–347.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.

Kelkar, Y.D., Strubczewski, N., Hile, S.E., Chiaromonte, F., Eckert, K.A., and Makova, K.D. (2010). What is a microsatellite: A computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. Genome Biol Evol 2, 620–635.

Keshtkar, E., Mathiassen, S.K., and Kudsk, P. (2017). No vegetative and fecundity fitness cost associated with acetyl-coenzyme a carboxylase non-target-site resistance in a black-grass (Alopecurus myosuroides huds) population. Front Plant Sci 8.

Kim, K.S., Kustu, S., and Inwood, W. (2006). Natural history of transposition in the green alga Chlamydomonas reinhardtii: Use of the AMT4 locus as an experimental system. Genetics 173, 2005–2019.

Knaus, B.J., and Grünwald, N.J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour 17, 44–53.

Koo, D.H., Molin, W.T., Saski, C.A., Jiang, J., Putta, K., Jugulam, M., Friebe, B., and Gill, B.S. (2018). Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed Amaranthus palmeri. Proceedings of the National Academy of Sciences 115, 3332–3337.

Koo, D.H., Sathishraj, R., Nakka, S., Ju, Y., Nandula, V.K., Jugulam, M., Friebe, B., and Gill, B.S. (2023a). Extrachromosomal circular DNA–mediated spread of herbicide resistance in interspecific hybrids of pigweed. Plant Physiol 193, 229–233.

Koo, D.H., Ju, Y., Putta, K., Sathishraj, R., Burgos, N., Jugulam, M., Friebe, B., and Gill, B.S. (2023b). Extrachromosomal DNA-mediated glyphosate resistance in Italian ryegrass. Pest Manag Sci 25–29.

Kopp, M., and Hermisson, J. (2007). Adaptation of a quantitative trait to a moving optimum. Genetics 176, 715–719.

Kopp, M., and Hermisson, J. (2009a). The genetic basis of phenotypic adaptation II: The distribution of adaptive substitutions in the moving optimum model. Genetics 183, 1453–1476.

Kopp, M., and Hermisson, J. (2009b). The genetic basis of phenotypic adaptation I: Fixation of beneficial mutations in the moving optimum model. Genetics 182, 233–249.

Koschwanez, J.H., Foster, K.R., and Murray, A.W. (2013). Improved use of a public good selects for the evolution of undifferentiated multicellularity. Elife 2013, 1–27.

Kosheleva, K., and Desai, M.M. (2018). Recombination Alters the Dynamics of Adaptation on Standing Variation in Laboratory Yeast Populations. Mol Biol Evol 35, 180–201.

Kouyos, R.D., Metcalf, C.J.E., Birger, R., Klein, E.Y., zur Wiesch, P.A., Ankomah, P., Arinaminpathy, N., Bogich, T.L., Bonhoeffer, S., Brower, C., et al. (2014). The path of least resistance: Aggressive or moderate treatment? Proc. R. Soc. B Biol. Sci. *281*.

Kreiner, J.M., Stinchcombe, J.R., and Wright, S.I. (2018). Population Genomics of Herbicide Resistance: Adaptation via Evolutionary Rescue. Annu Rev Plant Biol 69, 611–635.

Kronholm, I., and Collins, S. (2016). Epigenetic mutations can both help and hinder adaptive evolution. Mol Ecol 25, 1856–1868.

Kronholm, I., Bassett, A., Baulcombe, D., and Collins, S. (2017). Epigenetic and genetic contributions to adaptation in chlamydomonas. Mol Biol Evol 34, 2285–2306.

Kryazhimskiy, S., Tkăcik, G., and Plotkin, J.B. (2009). The dynamics of adaptation on correlated fitness landscapes. Proc Natl Acad Sci U S A 106, 18638–18643.

Lachapelle, J., and Bell, G. (2012). Evolutionary Rescue Of Sexual And Asexual Populations In A Deteriorating Environment. Evolution (N Y) 66, 3508–3518.

Lachapelle, J., Reid, J., and Colegrave, N. (2015a). Repeatability of adaptation in experimental populations of different sizes. Proc Biol Sci 282.

Lachapelle, J., Bell, G., and Colegrave, N. (2015b). Experimental adaptation to marine conditions by a freshwater alga. Evolution (N Y) 69, 2662–2675.

Lachapelle, J., Colegrave, N., and Bell, G. (2017a). The effect of selection history on extinction risk during severe environmental change. J Evol Biol 30, 1872–1883.

Laforest, M., Martin, S.L., Bisaillon, K., Soufiane, B., Meloche, S., and Page, E. (2020). A chromosome-scale draft sequence of the Canada fleabane genome. Pest Manag Sci 76, 2158–2169.

Lagator, M. (2012a). Experimental evolutionof herbicide resistance in Chlamydomonas reinhardtii-Chapter 7: General discussion. University of Warwick.

Lagator, M. (2012b). Experimental Evolution of Herbicide Resistance in Chlamydomonas reinhardtii.Chapter 6. Effects of herbicide dose - local adaptation along an environmental gradient. University of Warwick.

Lagator, M. (2012c). Experimental Evolution of Herbicide Resistance in Chlamydomonas reinhardtii Mato Lagator. University of Warwick.

Lagator, M., Vogwill, T., Mead, A., Colegrave, N., and Neve, P. (2013a). Herbicide mixtures at high doses slow the evolution of resistance in experimentally evolving populations of Chlamydomonas reinhardtii. New Phytol 198, 938–945.

Lagator, M., Vogwill, T., Colegrave, N., and Neve, P. (2013b). Herbicide cycling has diverse effects on evolution of resistance in Chlamydomonas reinhardtii. Evol Appl 6, 197–206.

Lagator, M., Colegrave, N., and Neve, P. (2014a). Selection history and epistatic interactions impact dynamics of adaptation to novel environmental stresses. Proc Biol Sci 281, 20141679.

Lagator, M., Morgan, A., Neve, P., and Colegrave, N. (2014b). Role of sex and migration in adaptation to sink environments. Evolution (N Y) 68, 2296–2305.

Lagator, M., Uecker, H., and Neve, P. (2021). Adaptation at different points along antibiotic concentration gradients. Biol Lett 17, 0–5.

Lang, G.I., Rice, D.P., Hickman, M.J., Sodergren, E., Weinstock, G.M., Botstein, D., and Desai, M.M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500, 571–574.

Larran, A.S., Palmieri, V.E., Tuesca, D., Permingeat, H.R., and Perotti, V.E. (2022). Coexistence of target-site and non-target-site mechanisms of glyphosate resistance in Amaranthus palmeri populations from Argentina. Acta Sci Agron 44, 1–8.

Lefebvre-Legendre, L., Choquet, Y., Kuras, R., Loubéry, S., Douchi, D., and Goldschmidt-Clermont, M. (2015). A nucleus-encoded chloroplast protein regulated by iron availability governs expression of the photosystem i subunit PsaA in Chlamydomonas reinhardtii. Plant Physiol 167, 1527–1540.

Lenski, R.E. (2017a). What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet 13.

Lenski, R.E. (2017b). Convergence and divergence in a long-term experiment with bacteria. American Naturalist 190, S57–S68.

Lenski, R.E., Rose, M.R., Simpson, S.C., and Tadler, S.C. (1991). Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. Am Nat 138, 1315–1341.

Levin, B.R., Perrot, V., and Walker, N. (2000). Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics 154, 985–997.

Levy, S.B., and Marshall, B. (2004). Antibacterial resistance worldwide : causes , challenges and responses REVIEW. 10, 122–129.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv preprint, 1–3.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Li, J., Peng, Q., Han, H., Nyporko, A., Kulynych, T., Yu, Q., and Powles, S.B. (2018). Glyphosate Resistance in Tridax procumbens via a Novel EPSPS Thr-102-Ser Substitution. J Agric Food Chem 66, 7880–7888.

Li, K., Xu, P., Wang, J., Yi, X., and Jiao, Y. (2023). Identification of errors in draft genome assemblies at single-nucleotide resolution for quality assessment and improvement. Nat Commun 14, 1–12.

Li, W., Guan, Q., Wang, Z.Y., Wang, Y., and Zhu, J. (2013). A bi-functional xyloglucan galactosyltransferase is an indispensable salt stress tolerance determinant in arabidop. Mol Plant 6, 1344–1354.

Lim, J., Lim, C.W., and Lee, S.C. (2020). Pepper Novel Serine-Threonine Kinase CaDIK1 Regulates Drought Tolerance via Modulating ABA Sensitivity. Front Plant Sci 11, 1–10.

Lindsey, H.A., Gallie, J., Taylor, S., and Kerr, B. (2013). Evolutionary rescue from extinction is contingent on a lower rate of environmental change. Nature 494, 463–467.

Liu, F., and Cao, Y. (2018). Expression of the 5-enoylpyruvylshikimate-3-phosphate synthase domain from the Acremonium sp. aroM complex enhances resistance to glyphosate. Biotechnol Lett 40, 855–864.

Liu, T., Arsenault, J., Vierling, E., and Kim, M. (2021). Mitochondrial ATP synthase subunit d, a component of the peripheral stalk, is essential for growth and heat stress tolerance in Arabidopsis thaliana. Plant Journal 107, 713–726.

Long, A., Liti, G., Luptak, A., and Tenaillon, O. (2015). Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat Rev Genet 16, 567–582.

Loubet, I., Meyer, L., Michel, S., Pernin, F., Carrère, S., Barrès, B., Le Corre, V., and Délye, C. (2023). A high diversity of non-target site resistance mechanisms to acetolactate-synthase (ALS) inhibiting herbicides has evolved within and among field populations of common ragweed (Ambrosia artemisiifolia L.). BMC Plant Biol 23, 1–16.

MacLean, R.C., and Buckling, A. (2009). The distribution of fitness effects of beneficial mutations in pseudomonas aeruginosa. PLoS Genet 5.

Majeed, Y., Zhu, X., Zhang, N., ul-Ain, N., Raza, A., Haider, F.U., and Si, H. (2023). Harnessing the role of mitogen-activated protein kinases against abiotic stresses in plants. Front Plant Sci 14, 1–15.

Martin, S.L., Benedict, L., Sauder, C.A., Wei, W., da Costa, L.O., Hall, L.M., and Beckie, H.J. (2017). Glyphosate resistance reduces kochia fitness: Comparison of segregating resistant and susceptible F2 populations. Plant Science 261, 69–79.

Matzrafi, M., Preston, C., and Brunharo, C.A. (2021). Review: evolutionary drivers of agricultural adaptation in Lolium spp. Pest Manag Sci 77, 2209–2218.

McDonald, M.J. (2019). Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery. EMBO Rep 20, 1–14.

Mei, Y., Xu, Y., Wang, S., Qiu, L., and Zheng, M. (2018). Investigation of glyphosate resistance levels and target-site based resistance (TSR) mechanisms in Conyza canadensis (L.) from apple orchards around areas of Bohai seas and Loess Plateau in China. Pestic Biochem Physiol 146, 7–12.

Melero-Jiménez, I.J., Bañares-España, E., Reul, A., Flores-Moya, A., and García-Sánchez, M.J. (2021). Detection of the maximum resistance to the herbicides diuron and glyphosate, and evaluation of its phenotypic cost, in freshwater phytoplankton. Aquatic Toxicology 240.

Menchari, Y., Chauvel, B., Darmency, H., and Délye, C. (2008). Fitness costs associated with three mutant acetylcoenzymeA carboxylase aeles endowing herbi resistance in black-gras Alopecurus myosuoides. Journal of Applied Ecology 45, 939–947.

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., et al. (2007). The Chlamydomonas genome Reveals the Evolution of Key Animal and Plant Functions. Science (1979) 318, 245–251.

Mesnage, R., Oestreicher, N., Poirier, F., Nicolas, V., Boursier, C., and Vélot, C. (2020). Transcriptome profiling of the fungus Aspergillus nidulans exposed to a commercial glyphosate-based herbicide under conditions of apparent herbicide tolerance. Environ Res 182, 109116.

Metcalf, W.W., and Wanner, B.L. (1993). Evidence for a fourteen-gene, phnC to phnP locus for phosphonate metabolism in Escherichia coli. Gene 129, 27–32.

Michitte, P., De Prado, R., Espinoza, N., Pedro Ruiz-Santaella, J., and Gauvrit, C. (2007). Mechanisms of Resistance to Glyphosate in a Ryegrass (Lolium Multiflorum) Biotype from Chile . Weed Sci 55, 435–440.

Molin, W.T., Wright, A.A., Lawton-Rauh, A., and Saski, C.A. (2017). The unique genomic landscape surrounding the EPSPS gene in glyphosate resistant Amaranthus palmeri: A repetitive path to resistance. BMC Genomics 18, 1–16.

Mora, A.D., Rosario, J., Rojano-Delgado, A.M., Palma-Bautista, C., Torra, J., Alcántara-De La Cruz, R., and De Prado, R. (2019). Multiple Resistance to Synthetic Auxin Herbicides and Glyphosate in Parthenium hysterophorus Occurring in Citrus Orchards. J Agric Food Chem 67, 10010–10017.

Moretti, M.L., Alárcon-Reverte, R., Pearce, S., Morran, S., and Hanson, B.D. (2017). Transcription of putative tonoplast transporters in response to glyphosate and paraquat stress in Conyza bonariensis and Conyza canadensis and selection of reference genes for qRT-PCR. PLoS One 12, 1–16.

Morgan, A.D., Ness, R.W., Keightley, P.D., and Colegrave, N. (2014). Spontaneous mutation accumulation in multiple strains of the green alga, Chlamydomonas reinhardtii. Evolution (N Y) 68, 2589–2602.

Morran, S., Moretti, M.L., Brunharo, C.A., Fischer, A.J., and Hanson, B.D. (2018). Multiple target site resistance to glyphosate in junglerice (Echinochloa colona) lines from California orchards. Pest Manag Sci 74, 2747–2753.

Muellner, J., and Schmidt, K.H. (2020). Yeast Genome Maintenance by the Multifunctional. Genes (Basel) 11, 6–8.

Murphy, Brent.P., and Tranel, P.J. (2019). Target-Site Mutations Conferring Herbicide Resistance. Plants 8.

Myers, J.P., Antoniou, M.N., Blumberg, B., Carroll, L., Colborn, T., Everett, L.G., Hansen, M., Landrigan, P.J., Lanphear, B.P., Mesnage, R., et al. (2016). Concerns over use of glyphosate-based herbicides and risks associated with exposures: A consensus statement. Environ Health 15, 1–13.

Nang, S.C., Morris, F.C., Mcdonald, M.J., Han, M., Wang, J., Strugnell, R.A., Velkov, T., and Li, J. (2018). Fitness cost of mcr-1 -mediated polymyxin resistance in Klebsiella pneumoniae. 1604–1610.

Ness, R.W., Morgan, A.D., Colegrave, N., and Keightley, P.D. (2012). Estimate of the spontaneous mutation rate in Chlamydomonas reinhardtii. Genetics 192, 1447–1454.

Ness, R.W., Morgan, A.D., Vasanthakrishnan, R.B., Colegrave, N., and Keightley, P.D. (2015). Extensive de novo mutation rate variation between individuals and across the genome of Chlamydomonas reinhardtii. Genome Res 25, 1739–1749.

Nestler, H., Groh, K.J., Schönenberger, R., Behra, R., Schirmer, K., Eggen, R.I.L.L., and Suter, M.J.-F. (2012a). Multiple-endpoint assay provides a detailed mechanistic view of responses to herbicide exposure in Chlamydomonas reinhardtii. Aquatic Toxicology 110–111, 214–224.

Nestler, H., Groh, K.J., Schönenberger, R., Eggen, R.I.L., and Suter, M.J.-F. (2012b). Linking proteome responses with physiological and biochemical effects in herbicide-exposed Chlamydomonas reinhardtii. J Proteomics 75, 5370–5385.

Neve, P., and Powles, S. (2005). Recurrent selection with reduced herbicide rates results in the rapid evolution of herbicide resistance in Lolium rigidum. Theoretical and Applied Genetics 110, 1154–1166.

Neve, P., Vila-Aiub, M., and Roux, F. (2009). Evolutionary-thinking in agricultural weed management. New Phytologist 184, 783–793.

Neve, P., Busi, R., Renton, M., and Vila-Aiub, M. (2014). Expanding the eco-evolutionary context of herbicide resistance research. Pest Manag Sci 70.

Newman, M.M., Hoilett, N., Lorenz, N., Dick, R.P., Liles, M.R., Ramsier, C., and Kloepper, J.W. (2016). Glyphosate effects on soil rhizosphere-associated bacterial communities. Sci. Total Environ. *543*, 155–160.

Ng, C.H., Wickneswari, R., Salmijah, S., Teng, Y.T., and Ismail, B.S. (2003). Gene polymorphisms in glyphosate-resistant and -susceptible biotypes of Eleusine indica from Malaysia. Weed Res 43, 108–115.

Nguyen, A.H., Matsui, A., Tanaka, M., Mizunashi, K., Nakaminami, K., Hayashi, M., Iida, K., Toyoda, T., Nguyen, D. Van, and Seki, M. (2014). Loss of Arabidopsis 5′-3′ Exoribonuclease AtXRN4 Function Enhances Heat Stress Tolerance of Plants Subjected to Severe Heat Stress. Plant Cell Physiol 56, 1762–1772.

Nguyen, H.M., Cuiné, S., Beyly-Adriano, A., Légeret, B., Billon, E., Auroy, P., Beisson, F., Peltier, G., and Li-Beisson, Y. (2013). The green microalga Chlamydomonas reinhardtii has a single ω-3 fatty acid desaturase that localizes to the chloroplast and impacts both plastidic and extraplastidic membrane lipids. Plant Physiol 163, 914–928.

Niu, Y., Zhang, Q., Wang, J., Li, Y., Wang, X., and Bao, Y. (2022). Vitamin E synthesis and response in plants. Front Plant Sci 13, 1–9.

O'Brien, T.F. (2002). Emergence, spread, and environmental effect of antimicrobial resistance: How use of an antimicrobial anywhere can increase resistance to any antimicrobial anywhere else. Clinical Infectious Diseases 34.

O'Donnell, S., Chaux, F., and Fischer, G. (2020). Highly Contiguous Nanopore Genome Assembly of CHlamydomonas reinhardtii CC-1690. Microbiology Resources Announcements 9, 6–7.

Oerke, E.C. (2006). Crop losses to pests. Journal of Agricultural Science 144, 31–43.

Osipitan, O.A., and Dille, J.A. (2017). Fitness outcomes related to glyphosate resistance in kochia (Kochia scoparia): What life history stage to examine? Front Plant Sci 8, 1–13.

Ouyang, S.Q., He, S.J., Liu, P., Zhang, W.K., Zhang, J.S., and Chen, S.Y. (2011). The role of tocopherol cyclase in salt stress tolerance of rice (Oryza sativa). Sci China Life Sci 54, 181–188.

Palma-Bautista, C., Vázquez-Garciá, J.G., Domínguez-Valenzuela, J.A., Ferreira Mendes, K., Alcántara De La Cruz, R., Torra, J., and De Prado, R. (2021). Non-Target-Site Resistance Mechanisms Endow Multiple Herbicide Resistance to Five Mechanisms of Action in Conyza bonariensis. J Agric Food Chem 69, 14792–14801.

Palma-Bautista, C., Vázquez-Garcia, J.G., López-Valencia, G., Domínguez-Valenzuela, J.A., Barro, F., and De Prado, R. (2023). Reduced Glyphosate Movement and Mutation of the EPSPS Gene (Pro106Ser) Endow Resistance in Conyza canadensis Harvested in Mexico. J Agric Food Chem 71, 4477–4487.

Pan, L., Yu, Q., Han, H., Mao, L.L., Nyporko, A., Fan, L., Bai, L., and Powles, S.B. (2019). Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in Echinochloa colona. Plant Physiol 181, pp.00979.2019.

Pan, L., Yu, Q., Wang, J., Han, H., Mao, L., Nyporko, A., Maguza, A., Fan, L., Bai, L., and Powles, S. (2021). An ABCC-type transporter endowing glyphosate resistance in plants. Proc Natl Acad Sci U S A 118, 1–11.

Patriarcheas, D., Momtareen, T., and Gallagher, J.E.G. (2023). Yeast of Eden: microbial resistance to glyphosate from a yeast perspective. Curr Genet.

Patterson, E.L., Pettinga, D.J., Ravet, K., Neve, P., and Gaines, T.A. (2018). Glyphosate resistance and EPSPS gene duplication: Convergent evolution in multiple plant species. Journal of Heredity 109, 117–125.

Payne, Z.L., Penny, G.M., Turner, T.N., and Dutcher, S.K. (2023). A gap-free genome assembly of Chlamydomonas reinhardtii and detection of translocations induced by CRISPR-mediated mutagenesis. Plant Commun 4, 100493.

Pedersen, B.P., Neve, P., Andreasen, C., and Powles, S.B. (2007). Ecological fitness of a glyphosate-resistant Lolium rigidum population: Growth and seed production along a competition gradient. Basic Appl Ecol 8, 258–268.

Pedotti, M., Rosini, E., Molla, G., Moschetti, T., Savino, C., Vallone, B., and Pollegioni, L. (2009). Glyphosate resistance by engineering the flavoenzyme glycine oxidase. Journal of Biological Chemistry 284, 36415–36423.

Penaloza-Vazquez, A., Mena, G.L., Herrera-Estrella, L., and Bailey, A.M. (1995). Cloning and sequencing of the genes involved in glyphosate utilization by Pseudomonas pseudomallei. Appl Environ Microbiol 61, 538–543.

Peng, Y., Abercrombie, L.L.G., Yuan, J.S., Riggins, C.W., Sammons, R.D., Tranel, P.J., and Stewart, C.N. (2010). Characterization of the horseweed (Conyza canadensis) transcriptome using GS-FLX 454

pyrosequencing and its application for expression analysis of candidate non-target herbicide resistance genes. Pest Manag Sci 66, 1053–1062.

Perotti, V.E., Larran, A.S., Palmieri, V.E., Martinatto, A.K., Alvarez, C.E., Tuesca, D., and 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.

Perron, G.G., Zasloff, M., and Bell, G. (2006). Experimental evolution of resistance to an antimicrobial peptide. Proceedings of the Royal Society B: Biological Sciences 273, 251–256.

Perron, G.G., Gonzalez, A., and Buckling, A. (2008). The rate of environmental change drives adaptation to an antibiotic sink. J Evol Biol 21, 1724–1731.

Petkovic, N., and Colegrave, N. (2023). The effects of sex on extinction dynamics of Chlamydomonas reinhardtii depend on the rate of environmental change. J Evol Biol 36, 1783–1795.

Petrey, A.C., and de la Motte, C.A. (2014). Hyaluronan, a crucial regulator of inflammation. Front Immunol 5, 1–13.

Piasecki, C., Carvalho, I.R., Cechin, J., Goulart, F.A.P., da Maia, L.C., Agostinetto, D., Caverzan, A., Stewart, C.N., and Vargas, L. (2019a). Oxidative stress and differential antioxidant enzyme activity in glyphosate-resistant and-sensitive hairy fleabane in response to glyphosate treatment. Bragantia 78, 379–396.

Piasecki, C., Yang, Y., Benemann, D.P., Kremer, F.S., Galli, V., Millwood, R.J., Cechin, J., Agostinetto, D., Maia, L.C., Vargas, L., et al. (2019b). Transcriptomic analysis identifies new non-target site glyphosate-resistance genes in conyza bonariensis. Plants 8.

Pollegioni, L., Schonbrunn, E., and Siehl, D. (2011). Molecular basis of glyphosate resistance - Different approaches through protein engineering. FEBS Journal 278, 2753–2766.

Powles, S.B. (2008). Evolved glyphosate-resistant weeds around the world: Lessons to be learnt. Pest Manag Sci 64, 360–365.

Powles, S.B., and Yu, Q. (2010). Evolution in action: Plants resistant to herbicides.

Powles, S.B., Lorraine-Colwill, D.F., and Dellow, J.J. (1998). Weed Science Society of America Evolved Resistance to Glyphosate in Rigid Ryegrass (Lolium rigidum) in. Weed Sci 46, 604–607.

Praat, M., De Smet, I., and Van Zanten, M. (2021). Protein kinase and phosphatase control of plant temperature responses. J Exp Bot 72, 7459–7473.

Pröschold, T., Harris, E.H., and Coleman, A.W. (2005). Portrait of a species: Chlamydonomas reinhardtii. Genetics 170, 1601–1610.

Purrington, C.B. (2000). Costs of resistance. Curr Opin Plant Biol 3, 305–308.

R Core Team (2023). R: A language and environment for statistical computing.

Rainey, P.B., and Travisano, M. (1998). Adaptive radiation in a Heterogeneous Environment. Nature 32, 69–72.

Ramón-Carbonell, M., López-Pérez, M., González-Candelas, L., and Sánchez-Torres, P. (2019). Pdmfs1 transporter contributes to penicilliun digitatum fungicide resistance and fungal virulence during citrus fruit infection. J. Fungi *5*

Randolph-Anderson, B.L., Sato, R., Johnson, A.M., Harris, E.H., Hauser, C.R., Oeda, K., Ishige, F., Nishio, S., Gillham, N.W., and Boynton, J.E. (1998). Isolation and characterization of a mutant protoporphyrinogen oxidase gene from Chlamydomonas reinhardtii conferring resistance to porphyric herbicides. Plant Mol Biol 38, 839–859.

Ravet, K., Patterson, E.L., Krähmer, H., Hamouzová, K., Fan, L., Jasieniuk, M., Lawton-Rauh, A., Malone, J.M., McElroy, J.S., Merotto, A., et al. (2018). The power and potential of genomics in weed biology and management. Pest Manag Sci 74, 2216–2225.

Ravishankar, A., Cumming, J.R., and Gallagher, J.E.G. (2020a). Mitochondrial metabolism is central for response and resistance of Saccharomyces cerevisiae to exposure to a glyphosate-based herbicide. Environmental Pollution 262, 114359.

Ravishankar, A., Pupo, A., and Gallagher, J.E.G. (2020b). Resistance mechanisms of saccharomyces cerevisiae to commercial formulations of glyphosate involve DNA damage repair, the cell cycle, and the cell wall structure. G3: Genes, Genomes, Genetics 10, 2043–2056.

Reboud, X. (2002). Response of Chlamydomonas reinhardtii to herbicides: Negative relationship between toxicity and water solubility across several herbicide families. Bull Environ Contam Toxicol 69, 554–561.

Reboud, X., and Bell, G. (1997). Experimental evolution in Chlamydomonas III. Evolution of specialist and generalist types in environments that vary in space and time. Heredity (Edinb) 78, 507–514.

Reboud, X., Majerus, N., Gasquez, J., and Powles, S. (2007). Chlamydomonas reinhardtii as a model system for pro-active herbicide resistance evolution research. Biological Journal of the Linnean Society 91, 257–266.

Reinbothe, S., Ortel, B., and Parthier, B. (1993). Overproduction by gene amplification of the multifunctional arom protein confers glyphosate tolerance to a plastid-free mutant of Euglena gracilis. MGG Molecular & General Genetics 239, 416–424.

Remy, E., Cabrito, T.R., Baster, P., Batista, R.A., Teixeira, M.C., Friml, J., Sá-Correia, I., and Duque, P. (2013). A Major Facilitator Superfamily transporter plays a dual role in polar auxin transport and drought stress tolerance in Arabidopsis. Plant Cell *25*, 901–926.

Richmond, M.E. (2018). Glyphosate: A review of its global use, environmental impact, and potential health effects on humans and other species. J Environ Stud Sci 8, 416–434.

Rong-Mullins, X., Ravishankar, A., McNeal, K.A., Lonergan, Z.R., Biega, A.C., Creamer, J.P., and Gallagher, J.E.G. (2017). Genetic variation in Dip5, an amino acid permease, and Pdr5, a multiple drug transporter, regulates glyphosate resistance in S. cerevisiae. PLoS One 12, 1–17.

Roux, F., and Reboud, X. (2005). Is the cost of herbicide resistance expressed in the breakdown of the relationships between characters? A case study using synthetic-auxin-resistant Arabidopsis thaliana mutants. Genet Res 85, 101–110.

Roux, F., Gasquez, J., and Reboud, X. (2004). The Dominance of the Herbicide Resistance Cost in Several Arabidopsis thaliana Mutant Lines. Genetics 166, 449–460.

Roux, F., Camilleri, C., Giancola, S., Brunel, D., and Reboud, X. (2005). Epistatic interactions among herbicide resistances in Arabidopsis thaliana: The fitness cost of multiresistance. Genetics 171, 1277–1288.

Rozen, D.E., Mcgee, L., Levin, B.R., and Klugman, K.P. (2007). Fitness Costs of Fluoroquinolone Resistance in Streptococcus pneumoniae . 51, 412–416.

Saier, M.H. (2000). A Functional-Phylogenetic Classification System for Transmembrane Solute Transporters. Microbiology and Molecular Biology Reviews 64, 354–411.

Sammons, R.D., and Gaines, T.A. (2014). Glyphosate resistance: State of knowledge. Pest Manag Sci 70, 1367–1377.

Sanchís, J., Kantiani, L., Llorca, M., Rubio, F., Ginebreda, A., Fraile, J., Garrido, T., and Farré, M. (2012). Erratum: Determination of glyphosate in groundwater samples using an ultrasensitive immunoassay and confirmation by on-line solid-phase extraction followed by liquid chromatography coupled to tandem mass spectrometry .Analytical and Bioanalytical Chemistr. Anal Bioanal Chem 404, 617.

Santos-Beneit, F. (2015). The Pho regulon: A huge regulatory network in bacteria. Front Microbiol 6, 1–13.

Saxton, M.A., Morrow, E.A., Bourbonniere, R.A., and Wilhelm, S.W. (2011). Glyphosate influence on phytoplankton community structure in Lake Erie. J. Great Lakes Res. *37*, 683–690.

Schenk, M.F., Szendro, I.G., Salverda, M.L.M., Krug, J., and De Visser, J.A.G.M. (2013). Patterns of epistasis between beneficial mutations in an antibiotic resistance gene. Mol Biol Evol 30, 1779–1787.

Schiffers, K., Bourne, E.C., Lavergne, S., Thuiller, W., and Travis, J.M.J. (2013). Limited evolutionary rescue of locally adapted populations facing climate change. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 0–9.

Schlötterer, C., Kofler, R., Versace, E., Tobler, R., and Franssen, S.U. (2015). Combining experimental evolution with next-generation sequencing: A powerful tool to study adaptation from standing genetic variation. Heredity (Edinb) 114, 431–440.

Schneider, T., Rizzardi, M.A., Grando, M.F., Ceccon, C.C., Nunes, A.L., and Didone, D. (2021). Expression of genes related to the mechanism of resistance of conyza sumatrensis to glyphosate. Acta Sci Agron 43, 1–8.

Schönbrunn, E., Eschenburg, S., Shuttleworth, W.A., Schloss, J. V, Amrhein, N., Evans, J.N., and 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, 1376–1380.

Schoustra, S., Hwang, S., Krug, J., and de Visser, J.A.G.M. (2016). Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. Proceedings of the Royal Society B: Biological Sciences 283.

Schwedt, I., Schöne, K., Eckert, M., Pizzinato, M., Winkler, L., Knotkova, B., Richts, B., Hau, J.L., Steuber, J., Mireles, R., et al. (2023). The low mutational flexibility of the EPSP synthase in Bacillus subtilis is due to a higher demand for shikimate pathway intermediates. Environ Microbiol 25, 3604–3622.

Sen, M.K., Hamouzová, K., Košnarová, P., Roy, A., and Soukup, J. (2022). Herbicide resistance in grass weeds: Epigenetic regulation matters too. Front Plant Sci 13, 1–9.

Shaner, D.L. (2009). Role of Translocation as A Mechanism of Resistance to Glyphosate. Weed Sci 57, 118–123.

Shaner, D.L. (2014). Lessons Learned From the History of Herbicide Resistance. Weed Sci 62, 427–431.

Shaner, D.L., Lindenmeyer, R.B., and Ostlie, M.H. (2012). What have the mechanisms of resistance to glyphosate taught us? Pest Manag Sci 68, 3–9.

Shinabarger, D.L., and Braymer, H.D. (1986). Glyphosate catabolism by Pseudomonas sp. strain PG2982. J Bacteriol 168, 702–707.

Shumate, A., and Salzberg, S.L. (2021). Liftoff: Accurate mapping of gene annotations. Bioinformatics 37, 1639–1643.

Shushkova, T. V., Vinokurova, N.G., Baskunov, B.P., Zelenkova, N.F., Sviridov, A. V., Ermakova, I.T., and Leontievsky, A.A. (2016). Glyphosate acetylation as a specific trait of Achromobacter sp. Kg 16 physiology. Appl Microbiol Biotechnol 100, 847–855.

Singh, K.P., Miaskowski, C., Dhruva, A.A., Flowers, E., and Kober, K.M. (2018). Mechanisms and Measurement of Changes in Gene Expression. Biol Res Nurs 20, 369–382.

Singh, S., Kumar, V., Gill, J.P.K., Datta, S., Singh, S., Dhaka, V., Kapoor, D., Wani, A.B., Dhanjal, D.S., Kumar, M., et al. (2020). Herbicide glyphosate: Toxicity and microbial degradation. Int J Environ Res Public Health 17, 1–18.

Slotboom, D.J., Konings, W.N., and Lolkema, J.S. (1999). Structural Features of the Glutamate Transporter Family. Microbiology and Molecular Biology Reviews 63, 293–307.

Soontharapirakkul, K., Promden, W., Yamada, N., Kageyama, H., Incharoensakdi, A., Iwamoto-Kihara, A., and Takabe, T. (2011). Halotolerant cyanobacterium Aphanothece halophytica contains an Na +-dependent F1F0-ATP synthase with a potential role in salt-stress tolerance. Journal of Biological Chemistry 286, 10169–10176.

Staub, J.M., Brand, L., Tran, M., Kong, Y., and Rogers, S.G. (2012). Bacterial glyphosate resistance conferred by overexpression of an E. coli membrane eZux transporter. J Ind Microbiol Biotechnol 39, 641–647.

Steinrucken, H.C., and Amrhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid-3-phosphate synthase. Biochem Biophys Res Commun 1207–1212.

Stosiek, N., Talma, M., and Klimek-Ochab, M. (2020). Carbon-Phosphorus Lyase—the State of the Art. Appl Biochem Biotechnol 190, 1525–1552.

Sun, Y., Zhao, J.Y., Li, Y.T., Zhang, P.G., Wang, S.P., Guo, J., Chen, J., Zhou, Y. Bin, Chen, M., Ma, Y.Z., et al. (2021). Genome-Wide Analysis of the C2 Domain Family in Soybean and Identification of a Putative Abiotic Stress Response Gene GmC2-148. Front Plant Sci 12, 1–19.

Sviridov, A. V., Shushkova, T. V., Ermakova, I.T., Ivanova, E. V., Epiktetov, D.O., and Leontievsky, A.A. (2015). Microbial degradation of glyphosate herbicides (review). Appl Biochem Microbiol 51, 188–195.

Takano, H.K., Fernandes, V.N.A., Adegas, F.S., Oliveira, R.S., Westra, P., Gaines, T.A., and Dayan, F.E. (2020). A novel TIPT double mutation in EPSPS conferring glyphosate resistance in tetraploid Bidens subalternans. Pest Manag Sci 76, 95–102.

Tani, E., Chachalis, D., and Travlos, I.S. (2015). A Glyphosate Resistance Mechanism in Conyza canadensis Involves Synchronization of EPSPS and ABC-transporter Genes. Plant Mol Biol Report 33, 1721–1730.

Tao, B., Shao, B.H., Qiao, Y.X., Wang, X.Q., Chang, S.J., and Qiu, L.J. (2017). Identification and functional analysis of a new glyphosate resistance gene from a fungus cDNA library. Pestic Biochem Physiol 140, 65–68.

Tardif, F.J., and Powles, S.B. (1994). Herbicide multiple-resistance in a Lolium rigidum biotype is endowed by multiple mechanisms: isolation of a subset with resistant acetyl-CoA carboxylase. Physiol Plant 91, 488–494.

Tenaillon, O., Toupance, B., Nagard, L., and Godelle, B. (1999). Mutators,populationsize,adaptive landscape and the adaptation of asexual populations of bacteria. Population (English Edition).

Thomas, P.D. (2017). The Gene Ontology and the meaning of biological function. In The Gene Ontology Handbook, C. Dessimoz, and N. Skunca, eds. (Springer Open), pp. 15–24.

Tolner, B., Ubbink-Kok, T., Poolman, B., and Konings, W.N. (1995). Characterization of the proton/glutamate symport protein of Bacillus subtilis and its functional expression in Escherichia coli. J Bacteriol 177, 2863–2869.

Toprak, E., Veres, A., Michel, J.B., Chait, R., Hartl, D.L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat Genet 44, 101–105.

Vallon, O., and Dutcher, S. (2008). Treasure hunting in the Chlamydomonas genome. Genetics 179, 3–6.

Vargas-Blanco, D.A., and Shell, S.S. (2020). Regulation of mRNA Stability During Bacterial Stress Responses. Front Microbiol 11.

Vemanna, R.S., Vennapusa, A.R., Easwaran, M., Chandrashekar, B.K., Rao, H., Ghanti, K., Sudhakar, C., Mysore, K.S., and Makarla, U. (2017). Aldo-keto reductase enzymes detoxify glyphosate and improve herbicide. Plant Biotechnol J 15, 794–804.

Vidi, P.A., Kanwischer, M., Baginsky, S., Austin, J.R., Csucs, G., Dörmann, P., Kessler, F., and Bréhélin, C. (2006). Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. Journal of Biological Chemistry 281, 11225–11234.

Vila-Aiub, M., Neve, P., and Powles, S.B. (2009). Fitness costs associated with evolved herbicide resistance alleles in plants. New Phytologist 184, 751–767.

Vila-Aiub, M., Balbi, M.C., Distéfano, A.J., Fernández, L., Hopp, E., Yu, Q., and Powles, S.B. (2011). Glyphosate resistance in perennial Sorghum halepense (Johnsongrass), endowed by reduced glyphosate translocation and leaf uptake. Pest Manag Sci 68, 430–436.

Vila-Aiub, M., Goh, S.S., Gaines, T.A., Han, H., Busi, R., Yu, Q., and Powles, S.B. (2014). No fitness cost of glyphosate resistance endowed by massive EPSPS gene amplification in Amaranthus palmeri. Planta 239, 793–801.

Vila-Aiub, M., Yu, Q., and Powles, S.B. (2019). Do plants pay a fitness cost to be resistant to glyphosate? New Phytologist 223, 532–547.

Vila-Aiub, M.M., Neve, P., and Powles, S.B. (2005). Resistance cost of a cytochrome P450 herbicide metabolism mechanism but not an ACCase target site mutation in a multiple resistant Lolium rigidum population. New Phytologist 167, 787–796.

Vila-Aiub, M.M., Neve, P., and Powles, S.B. (2009). Evidence for an ecological cost of enhanced herbicide metabolism in Lolium rigidum. Journal of Ecology 97, 772–780.

Vogwill, T., Lagator, M., Colegrave, N., and Neve, P. (2012). The experimental evolution of herbicide resistance in Chlamydomonas reinhardtii results in a positive correlation between fitness in the presence and absence of herbicides. J Evol Biol 25, 1955–1964.

Wan, Y., Wang, M., Chan, E.W.C., and Chen, S. (2021). Membrane Transporters of the Major Facilitator Superfamily Are Essential for Long-Term Maintenance of Phenotypic Tolerance to Multiple Antibiotics in E. coli. Microbiol. Spectr. *9*.

Wang, P., Hsu, C.C., Du, Y., Zhu, P., Zhao, C., Fu, X., Zhang, C., Paez, J.S., Macho, A.P., Andy Tao, W., et al. (2020). Mapping proteome-wide targets of protein kinases in plant stress responses. Proc Natl Acad Sci U S A 117, 3270–3280.

Wichman, H.A., Badgett, M.R., Scott, L.A., Boulianne, C.M., and Bull, J.J. (1999). Different trajectories of parallel evolution during viral adaptation. Science (1979) 285, 422–424.

Wicke, D., Schulz, L.M., Lentes, S., Scholz, P., Poehlein, A., Gibhardt, J., Daniel, R., Ischebeck, T., and Commichau, F.M. (2019). Identification of the first glyphosate transporter by genomic adaptation. Environ Microbiol 21, 1287–1305.

Wilke, C.O. (2004). The speed of adaptation in large asexual populations. Genetics 167, 2045–2053.

Willi, Y., Van Buskirk, J., and Hoffmann, A.A. (2006). Limits to the adaptive potential of small populations. Annu Rev Ecol Evol Syst 37, 433–458.

Williams, G.M., Kroes, R., and Munro, I.C. (2000). Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. Regulatory Toxicology and Pharmacology 31, 117–165.

Wirth, M.A., Schulz-Bull, D.E., and Kanwischer, M. (2021). The challenge of detecting the herbicide glyphosate and its metabolite AMPA in seawater – Method development and application in the Baltic Sea. Chemosphere 262, 128327.

Wiser, M.J., Ribeck, N., and Lenski, R.E. (2013). Long-term dynamics of adaptationn asexual populations. Science (1979) 342, 1364–1367.

Yamamoto, H., Tobisawa, Y., Inubushi, T., Irie, F., Ohyama, C., and Yamaguchi, Y. (2017). A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. Journal of Biological Chemistry 292, 7304–7313.

Yanniccari, M., Vila-Aiub, M., Istilart, C., Acciaresi, H., and Castro, A.M. (2016). Glyphosate Resistance in Perennial Ryegrass (Lolium perenne L.) is Associated with a Fitness Penalty. 71–79.

Yanniccari, M., Gómez-Lobato, M.E., Istilart, C., Natalucci, C., Giménez, D.O., and Castro, A.M. (2017). Mechanism of resistance to glyphosate in Lolium perenne from Argentina. Front Ecol Evol 5, 1–8.

Ye, Q., Yu, J., Zhang, Z., Hou, L., and Liu, X. (2020). VvBAP1, a Grape C2 Domain Protein, Plays a Positive Regulatory Role Under Heat Stress. Front Plant Sci 11, 1–12.

Yu, Q., Han, H., Vila-Aiub, M.M., and Powles, S.B. (2010). AHAS herbicide resistance endowing mutations: Effect on AHAS functionality and plant growth. J Exp Bot 61, 3925–3934.

Yuan, J.S., Tranel, P.J., and Stewart, C.N. (2007). Non-target-site herbicide resistance: a family business. Trends Plant Sci 12, 6–13.

Zaprasis, A., Bleisteiner, M., Kerres, A., Hoffmann, T., and Bremer, E. (2015). Uptake of amino acids and their metabolic conversion into the compatible solute proline confers osmoprotection to Bacillus subtilis. Appl Environ Microbiol 81, 250–259.

Zhang, C. (2023). Subtelomeric 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) copy number variation confers glyphosate resistance in Eleusine indica. Res Sq 0–24.

Zhang, C., Feng, L., He, T.T., Yang, C.H., Chen, G.Q., and Tian, X.S. (2015). Investigating the mechanisms of glyphosate resistance in goosegrass (Eleusine indica) population from South China. J Integr Agric 14, 909–918.

Zhang, F., Zeng, D., Huang, L., Shi, Y., Chen, T., Zhang, F., and Zhou, Y. (2019). Stress-Activated Protein Kinase OsSAPK9 Regulates Tolerance to Salt Stress and Resistance to Bacterial Blight in Rice. Rice 12.

Zhang, Y., Chen, Z., Li, X., Wu, X., Chen, L., and Wang, G. (2023). Photosynthesis Responses of Tibetan Freshwater Algae Chlorella vulgaris to Herbicide Glyphosate. Int. J. Environ. Res. Public Health *20*.

Zhang, Y., Held, M.A., and Showalter, A.M. (2020). Elucidating the roles of three β-glucuronosyltransferases (GLCATs) acting on arabinogalactan-proteins using a CRISPR-Cas9 multiplexing approach in Arabidopsis. BMC Plant Biol 20.

Zhang, Z., Zhao, C., Wu, T., Xu, Y., Wang, L., and Niu, Y. (2024). Physiological response and molecular mechanism of glyphosate degradation by Pseudomonas alcaligenes Z1–1. Biochem Eng J 209, 109405.

Zhao, C., Zayed, O., Zeng, F., Liu, C., Zhang, L., Zhu, P., Hsu, C.C., Tuncil, Y.E., Tao, W.A., Carpita, N.C., et al. (2019). Arabinose biosynthesis is critical for salt stress tolerance in Arabidopsis. New Phytologist 224, 274–290.

Zhao, H., Tao, K., Zhu, J., Liu, S., Gao, H., and Zhou, X. (2015). Bioremediation potential of glyphosate-degrading pseudomonas spp. Strains isolated from contaminated soil. Journal of General and Applied Microbiology 61, 165–170.

Zhou, F.Y., Han, H., Han, Y.J., Nyporko, A., Yu, Q., Beckie, H.J., and Powles, S.B. (2023). Aldo-keto reductase may contribute to glyphosate resistance in Lolium rigidum. Pest Manag Sci 79, 1528–1537.

Zhou, K., Aertsen, A., and Michiels, C.W. (2014). The role of variable DNA tandem repeats in bacterial adaptation. FEMS Microbiol Rev 38, 119–141.

Zuo, S., Yi, Y., Wang, C., Li, X., Zhou, M., Peng, Q., Zhou, J., Yang, Y., and He, Q. (2022). Extrachromosomal Circular DNA (eccDNA): From Chaos to Function. Front Cell Dev Biol 9, 1–8.

Appendix 1 Glyphosate resistance mechanisms in weeds

Table reporting known molecular basis of glyphosate resistance per resistance mechanisms, weed species in which they have been reported and methods used. Reference list is found below the table.

Mechanisms	Species	Method
Target site		
EPSPS Mutations		
Pro15Thr	*Digitaria insularis*[1]	Sequencing [1]
Pro106Ser	*Eleusine indica*[2–8], *Amaranthus tuberculatus*[9,10], *Lolium rigidum*[11–14], *Lolium multiflorum*[15–17], *Echinochloa colona*[18,19], *Bidens pilosa*[20], *Parthenium hysterophorus*[21,22], *Chloris virgata*[23], *Conyza manicar*[24,25], *Amaranthus palmeri*[26], *Lolium perenne*[27], *Ambrosia artemisiifolia*[28], *Ambrosia trifida*[29], *Conyza canadensis*[30], *Carduus acanthoides* [31].	RT-PCR[2,7], Sequencing[3,8,9,11,13,14,16–19,21,23–25,27–32], RFLP and PCR-RFLP[5] ,PASA analysis[4,6], cloning[10,12,15,20,22,26].
Pro106Ala	*L. rigidum*[33], *L. multiflorum*[17], *E. indica*[34]	Sequencing[17,33,34]
Pro106Thr	*E. indica*[5,6], *L. rigidum*[11,35], *E. colona*[18,36,37], *Conyza sumatrensis*[38]	RT-PCR[5,6], Sequencing[11,18,21,35–38]
Pro106Leu	*L. rigidum*[12], *C. virgata*[23], *E. indica*[32]	Cloning[12] , Sequencing[23]
Pro182Thr	*C. sumatrensis*[39]	Sequencing[39]
Thr102Ser	*Tridax procumbens*[40]	Sequencing[40]
Tyr143Cys	*D. insularis*[1]	Sequencing[1]
Pro182Thr Tyr310Cys	*D.insularis*[41]	Sequencing[41]

Thr102Ile Pro106Ser (TIPS mutation)	*E. indica*[34,42,43], *B. pilosa*[20], *P. hysterophorus*[44]	Cloning[20], Sequencing[34,44] and dCAPS[42,43]
Thr102Ile Pro106Thr (TIPT mutation)	*Bidens subalternans*[45], *Euphorbia heterophylla*[46]	Cloning[45,46]
Pro106Thr Pro106Leu (polyploids)	*E. colona*[18,37]	Sequencing[18,37]
Thr102Ile Ala103Val Pro106Ser (TIAVPS mutation)	*Amaranthus hybridus*[47,48,49]	Sequencing[47–49]
EPSPS Gene Copy Number variation		
EPSPS Cassette	*A.palmeri*[50–52]	Sequencing, mapping[51] PCR[50,52] and gene expression[52]
EPSPS eccDNA	*A.palmeri*[53,54], *Amaranthus spinosus*[55], *L. perenne ssp multiflorum*[56]	Cloning[53 54 55], Fiber-FISH[53 54 55 56], Sequencing and Transcriptional activity[53], PFGE and DNA blot[56]
Subtelomeric CNV	*E.indica*[57]	WGSeq and assembly of GR and GS individuals
Tandem duplication	*Koshia scoparia* [58]	FISH and Fiber-FISH[58]
Unknown	*E. indica*[32],*Helianthus annuus*[59],*Helianthus glaucum*[60], *Chloris truncata*[61], *Bromus diandrus*[62], *A. palmeri*[63–65], *L. multiflorum*[66], *Poa annua*[67], *A. tuberculatus*[68]	RT-PCR of RNA[32], qPCR of Genomic DNA[59–61,63–68], testing absorption and translocation, shikimate assay, Sequencing[61]

EPSPS gene increased transcription	*L. perenne*[69], *E. indica*[34,70], *C. canadensis*[71]	RT PCR[34,69–71]
EPSPS increased enzyme activity	*D. insularis*[1]	Bradford[1]
Non-Target Site		
Reduced uptake		
Unknown	*Sorghum halepense*[72], *L. multiflorum*[73], *Conyza bonariensis*[74], *P. hysterophorus*[44]	14C-Glyphosate[44,72–74]
Reduced Translocation		
Sequestration in vacuole		
ABC transporters	*C. canadensis*[75,76], *C. bonariaensis*[77], *C. sumatrensis*[78], *Lolium ssp* [79]	Transcriptome sequencing[75] and RT-PCR[75–79]
Unknown	*S. halepense*[72], *C. canadensis*[30], *C. bonariensis*[74], *P. hysterophorus*[44]	14C-Glyphosate[30,44,72,74]
Transport outside of cytoplasm		
ABC transporters	*E. colona* – EcABCC8 gene[80]	RNAseq and RT-PCR[80]
Unknown	*S. halepense*[81]	EPSPS activity, EPSPS gene sequencing, glyphosate metabolism, 14C-Glyphosate[81]

Enhanced metabolic degradation		
Aldo-Keto reductase (AKR)	*E. colona* (EcAKR4-1)[82], *L. rigidum* (LrAKR4C10 and LrAKR1)[83], *E. indicia* (AKR4C10)[34]	RNA seq for gene discovery[34,82], RTqPCR[34], cDNA Sequencing[83], *Oryza sativa*[82] and E.coli[83] transformation
Detoxification		
Cytochrome P450s	*Ipomoea purpurea*[84,85], *E. indicia(CYP88)*[34], *C. canadensis*[86], *C. bonariensis*[87]	RNA seq for gene discovery[84,87],RTqPCR[34,87], Genome Wide outlier screen[85], Exome resequencing[85] or Genome sequencing[86] and GO[85,86]
Glycosyltransferase	*I. purpurea*[85], *C. canadensis*[86], *C. bonariensis*[87]	RNA seq for gene discovery and RTqPCR[87], Genome Wide outlier screen[85], Exome resequencing[85] or Genome sequencing[86] and GO[85,86]
Glutathione S-transferase (GST)	*I. purpurea*[85], *C. bonariensis*[87]	RNA seq for gene discovery and RTqPCR[87], Genome Wide outlier screen, Exome resequencing and GO[85]
ABC transporters	*I. purpurea*[85], *E. indica* (ABCC4)[34], *C. canadensis*[86], *C. bonariensis*[87]	RNA seq for gene discovery[84,87],RTqPCR[34,87], Genome Wide outlier screen[85], Exome resequencing[85] or Genome sequencing[86] and GO[85,86]
Catalase (CAT)	*C. bonariensis*[87,88]	RNA seq for gene discovery and RTqPCR[87], Enzymatic activity[88]

Peroxidase (POD)	*C. bonariensis*[87]	RNA seq for gene discovery and RTqPCR[87]
Superoxide dismutase (SOD)	*C. bonariensis*[87,88]	RNA seq for gene discovery and RTqPCR[87], Enzymatic activity[88]
Ascorbate peroxidase (APX)	*C. bonariensis*[88], *A. trifida (Harre et al, 2018)*	Enzymatic activity[88,89]
Glutathione reductase	*A. trifida (Harre et al, 2018)*	Enzymatic activity[89]
Dehydroascorbate reductase	*A. trifida (Harre et al, 2018)*	Enzymatic activity[89]

Reference list for Appendix1 :

1. Galeano, E., Barroso, A.A.M., Vasconcelos, T.S., López-Rubio, A., Albrecht, A.J.P., Victoria Filho, R., and Carrer, H. (2016). EPSPS variability, gene expression, and enzymatic activity in glyphosate-resistant biotypes of Digitaria insularis. Genet. Mol. Res. *15*, 1–15. 10.4238/gmr.15038730.
2. Baerson, S.R., Rodriguez, D.J., Tran, M., Feng, Y., Biest, N.A., and Dill, G.M. (2002). Glyphosate-resistant goosegrass. Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. Plant Physiol. *129*, 1265–1275. 10.1104/pp.001560.
3. Janel, L.H., Riggins, C.W., Steckel, L.E., and Tranel, P.J. (2016). The EPSPS Pro106Ser substitution solely accounts for glyphosate resistance in a goosegrass (Eleusine indica) population from Tennessee, United States. J. Integr. Agric. *15*. 10.1016/S2095-3119(15)61220-5.
4. Kaundun, S.S., Zelaya, I.A., Dale, R.P., Lycett, A.J., Carter, P., Sharples, K.R., and McIndoe, E. (2008). Importance of the P106S Target-Site Mutation in Conferring Resistance to Glyphosate in a Goosegrass (Eleusine indica) Population from the Philippines . Weed Sci. *56*, 637–646. 10.1614/ws-07-148.1.
5. Ng, C.H., Wickneswari, R., Salmijah, S., Teng, Y.T., and Ismail, B.S. (2003). Gene polymorphisms in glyphosate-resistant and -susceptible biotypes of Eleusine indica from Malaysia. Weed Res. *43*, 108–115. 10.1046/j.1365-3180.2003.00322.x.
6. Ng, C.H., Wickneswary, R., Salmijah, S., Teng, Y.T., and Ismail, B.S. (2004). Glyphosate resistance in Eleusine indica (L.) Gaertn. from different origins and polymerase chain reaction amplification of specific alleles. Aust. J. Agric. Res. *55*, 407–414. 10.1071/AR03155.
7. Takano, H.K., Mendes, R.R., Scoz, L.B., Lopez Ovejero, R.F., Constantin, J., Gaines, T.A., Westra, P., Dayan, F.E., and Oliveira, R.S. (2019). Proline-106 EPSPS mutation imparting glyphosate resistance in goosegrass (Eleusine indica) Emerges in South America. Weed

Sci. *67*, 48–56. 10.1017/wsc.2018.71.

8. Vázquez-García, J.G., Alcántara-De La Cruz, R., Rojano-Delgado, A.M., Palma-Bautista, C., De Portugal Vasconcelos, J.M., and De Prado, R. (2021). Multiple Herbicide Resistance Evolution: The Case of Eleusine indica in Brazil. J. Agric. Food Chem. *69*, 1197–1205. 10.1021/acs.jafc.0c03999.
9. Bell, M.S., Hager, A.G., and Tranel, P.J. (2013). Multiple Resistance to Herbicides from Four Site-of-Action Groups in Waterhemp (Amaranthus tuberculatus) . Weed Sci. *61*, 460–468. 10.1614/ws-d-12-00166.1.
10. Nandula, V.K., Ray, J.D., Ribeiro, D.N., Pan, Z., and Reddy, K.N. (2013). Glyphosate Resistance in Tall Waterhemp (Amaranthus tuberculatus) from Mississippi is due to both Altered Target-Site and Nontarget-Site Mechanisms . Weed Sci. *61*, 374–383. 10.1614/ws-d-12-00155.1.
11. Bostamam, Y., Malone, J.M., Dolman, F.C., Boutsalis, P., and Preston, C. (2012). Rigid Ryegrass (Lolium rigidum) Populations Containing a Target Site Mutation in EPSPS and Reduced Glyphosate Translocation Are More Resistant to Glyphosate . Weed Sci. *60*, 474–479. 10.1614/ws-d-11-00154.1.
12. Collavo, A., and Sattin, M. (2012). Resistance to glyphosate in Lolium rigidum selected in Italian perennial crops: Bioevaluation, management and molecular bases of target-site resistance. Weed Res. *52*, 16–24. 10.1111/j.1365-3180.2011.00883.x.
13. Fernández-Moreno, P.T., Travlos, I., Brants, I., and De Prado, R. (2017). Different levels of glyphosate-resistant Lolium rigidum L. among major crops in southern Spain and France. Sci. Rep. *7*, 2–13. 10.1038/s41598-017-13384-2.
14. Simarmata, M., and Penner, D. (2008). The Basis for Glyphosate Resistance in Rigid Ryegrass (Lolium Rigidum) from California. Weed Sci. *56*, 181–188. 10.1614/ws-07-057.1.
15. Perez-Jones, A., Park, K.W., Polge, N., Colquhoun, J., and Mallory-Smith, C.A. (2007). Investigating the mechanisms of glyphosate resistance in Lolium multiflorum. Planta *226*, 395–404. 10.1007/s00425-007-0490-6.
16. González-Torralva, F., Gil-Humanes, J., Barro, F., Brants, I., and De Prado, R. (2012). Target site mutation and reduced translocation are present in a glyphosate-resistant Lolium multiflorum Lam. biotype from Spain. Plant Physiol. Biochem. *58*, 16–22. 10.1016/j.plaphy.2012.06.001.
17. Jasieniuk, M., Ahmad, R., Sherwood, A.M., Firestone, J.L., Perez-Jones, A., Lanini, W.T., Mallory-Smith, C., and Stednick, Z. (2008). Glyphosate-Resistant Italian Ryegrass (Lolium multiflorum) in California: Distribution, Response to Glyphosate, and Molecular Evidence for an Altered Target Enzyme . Weed Sci. *56*, 496–502. 10.1614/ws-08-020.1.
18. Morran, S., Moretti, M.L., Brunharo, C.A., Fischer, A.J., and Hanson, B.D. (2018). Multiple target site resistance to glyphosate in junglerice (Echinochloa colona) lines from California orchards. Pest Manag. Sci. *74*, 2747–2753. 10.1002/ps.5061.
19. Alarcón-Reverte, R., García, A., Urzúa, J., and Fischer, A.J. (2013). Resistance to Glyphosate in Junglerice (Echinochloa colona) from California . Weed Sci. *61*, 48–54. 10.1614/ws-d-12-00073.1.
20. 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., and 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*, 1–12. 10.3389/fpls.2016.01492.

21. Palma-Bautista, C., Gherekhloo, J., Domínguez-Martínez, P.A., Domínguez-Valenzuela, J.A., Cruz-Hipolito, H.E., Alcántara-de la Cruz, R., Rojano-Delgado, A.M., and De Prado, R. (2019). Characterization of three glyphosate resistant Parthenium hysterophorus populations collected in citrus groves from Mexico. Pestic. Biochem. Physiol. *155*, 1–7. 10.1016/j.pestbp.2018.11.002.
22. Bracamonte, E., Fernández-Moreno, P.T., Barro, F., and de Prado, R. (2016). Glyphosate-resistant Parthenium hysterophorus in the Caribbean Islands: Non target site resistance and target site resistance in relation to resistance levels. Front. Plant Sci. *7*. 10.3389/fpls.2016.01845.
23. Ngo, T.D., Krishnan, M., Boutsalis, P., Gill, G., and Preston, C. (2018). Target-site mutations conferring resistance to glyphosate in feathertop Rhodes grass (Chloris virgata) populations in Australia. Pest Manag. Sci. *74*, 1094–1100. 10.1002/ps.4512.
24. Beres, Z.T., Giese, L.A., Mackey, D.M., Owen, M.D.K., Page, E.R., and Snow, A.A. (2020). Target-site EPSPS Pro-106-Ser mutation in Conyza canadensis biotypes with extreme resistance to glyphosate in Ohio and Iowa, USA. Sci. Rep. *10*, 1–9. 10.1038/s41598-020-64458-7.
25. Page, E.R., Grainger, C.M., Laforest, M., Nurse, R.E., Rajcan, I., Bae, J., and Tardif, F.J. (2018). Target and Non-target site Mechanisms Confer Resistance to Glyphosate in Canadian Accessions of Conyza canadensis. Weed Sci. *66*, 234–245. 10.1017/wsc.2017.69.
26. Dominguez-Valenzuela, J.A., Gherekhloo, J., Fernández-Moreno, P.T., Cruz-Hipolito, H.E., Alcántara-de la Cruz, R., Sánchez-González, E., and De Prado, R. (2017). First confirmation and characterization of target and non-target site resistance to glyphosate in Palmer amaranth (Amaranthus palmeri) from Mexico. Plant Physiol. Biochem. *115*, 212–218. 10.1016/j.plaphy.2017.03.022.
27. Ghanizadeh, H., Harrington, K.C., James, T.K., Woolley, D.J., and Ellison, N.W. (2015). Mechanisms of glyphosate resistance in two perennial ryegrass (Lolium perenne) populations. Pest Manag. Sci. *71*, 1617–1622. 10.1002/ps.3968.
28. Laforest, M., Martin, S.L., Bisaillon, K., Soufiane, B., Meloche, S., Tardif, F.J., and Page, E. (2024). The ancestral karyotype of the Heliantheae Alliance, herbicide resistance, and human allergens: Insights from the genomes of common and giant ragweed. Plant Genome *17*, 1–28. 10.1002/tpg2.20442.
29. Page, E.R., Martin, S., Meloche, S., Thibodeau, A., Nurse, R.E., Sikkema, P.H., Tardif, F.J., Cowbrough, M.J., and Laforest, M. (2024). Revisiting the origins of glyphosate-resistant giant ragweed (Ambrosia trifida L.) in Canada. Can. J. Plant Sci. *104*, 124–128. 10.1139/cjps-2023-0145.
30. Palma-Bautista, C., Vázquez-Garcia, J.G., López-Valencia, G., Domínguez-Valenzuela, J.A., Barro, F., and De Prado, R. (2023). Reduced Glyphosate Movement and Mutation of the EPSPS Gene (Pro106Ser) Endow Resistance in Conyza canadensis Harvested in Mexico. J. Agric. Food Chem. *71*, 4477–4487. 10.1021/acs.jafc.2c07833.
31. Palma-Bautista, C., Belluccini, P., Vázquez-García, J.G., Alcántara-de la Cruz, R., Barro, F., Portugal, J., and De Prado, R. (2023). Target-site and non-target-site resistance mechanisms confer multiple resistance to glyphosate and 2,4-D in Carduus acanthoides. Pestic. Biochem. Physiol. *191*. 10.1016/j.pestbp.2023.105371.
32. Chen, J., Li, Z., Cui, H., Yu, H., and Li, X. (2023). Gene Amplification of EPSPS with a Mutation in Conserved Region: The Evolved Glyphosate Resistance Mechanism in Eleusine indica. Agronomy *13*. 10.3390/agronomy13030699.
33. Yu, Q., Cairns, A., and Powles, S.B. (2007). Glyphosate, paraquat and ACCase multiple herbicide resistance evolved in a Lolium rigidum

biotype. Planta *225*, 499–513. 10.1007/s00425-006-0364-3.

34. Deng, W., Duan, Z., Li, Y., Peng, C., and Yuan, S. (2022). Multiple Resistance Mechanisms Involved in Glyphosate Resistance in Eleusine indica. Plants *11*. 10.3390/plants11233199.
35. Wakelin, A.M., and Preston, C. (2006). A target-site mutation is present in a glyphosate-resistant Lolium rigidum population. Weed Res. *46*, 432–440. 10.1111/j.1365-3180.2006.00527.x.
36. Alarcón-Reverte, R., García, A., Watson, S.B., Abdallah, I., Sabaté, S., Hernández, M.J., Dayan, F.E., and Fischer, A.J. (2015). Concerted action of target-site mutations and high EPSPS activity in glyphosate-resistant junglerice (Echinochloa colona) from California. Pest Manag. Sci. *71*, 996–1007. 10.1002/ps.3878.
37. Han, H., Yu, Q., Widderick, M.J., and Powles, S.B. (2016). Target-site EPSPS Pro-106 mutations: Sufficient to endow glyphosate resistance in polyploid Echinochloa colona? Pest Manag. Sci. *72*, 264–271. 10.1002/ps.4038.
38. Amaro-Blanco, I., Fernández-Moreno, P.T., Osuna-Ruiz, M.D., Bastidac, F., and Prado, R. De (2018). Mechanisms of glyphosate resistance and response to alternative herbicide-based management in populations of the three Conyza species introduced in southern Spain.pdf. Pest Manag. Sci. *74*, 1925–1937.
39. González-Torralva, F., Gil-Humanes, J., Barro, F., Domínguez-Valenzuela, J.A., and De Prado, R. (2014). First evidence for a target site mutation in the EPSPS2 gene in glyphosate-resistant Sumatran fleabane from citrus orchards. Agron. Sustain. Dev. *34*, 553–560. 10.1007/s13593-013-0163-8.
40. Li, J., Peng, Q., Han, H., Nyporko, A., Kulynych, T., Yu, Q., and Powles, S.B. (2018). Glyphosate Resistance in Tridax procumbens via a Novel EPSPS Thr-102-Ser Substitution. J. Agric. Food Chem. *66*, 7880–7888. 10.1021/acs.jafc.8b01651.
41. De Carvalho, L.B., Alves, P.L.D.C.A., González-Torralva, F., Cruz-Hipolito, H.E., Rojano-Delgado, A.M., De Prado, R., Gil-Humanes, J., Barro, F., and Luque De Castro, M.D. (2012). Pool of resistance mechanisms to glyphosate in digitaria insularis. J. Agric. Food Chem. *60*, 615–622. 10.1021/jf204089d.
42. Han, H., Vila-aiub, M., Jalaludin, A., Yu, Q., and Powles, S.B. (2017). A double EPSPS gene mutation endowing glyphosate resistance shows a remarkably high resistance cost. Plant Cell Environ. *40*. 10.1111/pce.13067.
43. Yu, Q., Jalaludin, A., Han, H., Chen, M., Douglas Sammons, R., and Powles, S.B. (2015). Evolution of a double amino acid substitution in the 5-enolpyruvylshikimate-3-phosphate synthase in Eleusine indica conferring high-level glyphosate resistance. Plant Physiol. *167*, 1440–1447. 10.1104/pp.15.00146.
44. Mora, A.D., Rosario, J., Rojano-Delgado, A.M., Palma-Bautista, C., Torra, J., Alcántara-De La Cruz, R., and De Prado, R. (2019). Multiple Resistance to Synthetic Auxin Herbicides and Glyphosate in Parthenium hysterophorus Occurring in Citrus Orchards. J. Agric. Food Chem. *67*, 10010–10017. 10.1021/acs.jafc.9b03988.
45. Takano, H.K., Fernandes, V.N.A., Adegas, F.S., Oliveira, R.S., Westra, P., Gaines, T.A., and Dayan, F.E. (2020). A novel TIPT double mutation in EPSPS conferring glyphosate resistance in tetraploid Bidens subalternans. Pest Manag. Sci. *76*, 95–102. 10.1002/ps.5535.
46. Mendes, R.R., Takano, H.K., Leal, J.F., Souza, A.S., Morran, S., Oliveira, R.S., Adegas, F.S., Gaines, T.A., and Dayan, F. (2020). Evolution of EPSPS double mutation imparting glyphosate resistance in wild poinsettia (Euphorbia heterophylla L.). PLoS One *15*, 1–17.

10.1371/journal.pone.0238818.

47. Perotti, V.E., Larran, A.S., Palmieri, V.E., Martinatto, A.K., Alvarez, C.E., Tuesca, D., and 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. 10.1002/ps.5303.
48. García, M.J., Palma-Bautista, C., Rojano-Delgado, A.M., Bracamonte, E., Portugal, J., Alcántara-De la Cruz, R., and 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*. 10.3390/ijms20102396.
49. Shabangu, N.M., Simelane, N.A., Reinhardt, C.F., Marais, D., and Vorster, B.J. (2024). First report of target-site resistance to glyphosate in Amaranthus hybridus L. in the Republic of South Africa. Weed Res. *64*, 290–301. 10.1111/wre.12635.
50. Manicardi, A., Milani, A., Scarabel, L., Mora, G., Recasens, J., Llenes, J.M., Montull, J.M., and Torra, J. (2023). First report of glyphosate resistance in an Amaranthus palmeri population from Europe. Weed Res., 1–6. 10.1111/wre.12579.
51. Molin, W.T., Wright, A.A., Lawton-Rauh, A., and Saski, C.A. (2017). The unique genomic landscape surrounding the EPSPS gene in glyphosate resistant Amaranthus palmeri: A repetitive path to resistance. BMC Genomics *18*, 1–16. 10.1186/s12864-016-3336-4.
52. Molin, W.T., Wright, A.A., VanGessel, M.J., McCloskey, W.B., Jugulam, M., and Hoagland, R.E. (2017). Survey of the genomic landscape surrounding the 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene in glyphosate-resistant Amaranthus palmeri from geographically distant populations in the USA.pdf. Pest Manag. Sci. *74*, 1109–1117.
53. Koo, D.H., Molin, W.T., Saski, C.A., Jiang, J., Putta, K., Jugulam, M., Friebe, B., and Gill, B.S. (2018). Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed Amaranthus palmeri. Proc. Natl. Acad. Sci. *115*, 3332–3337. 10.1073/pnas.1719354115.
54. Koo, D.H., Sathishraj, R., Friebe, B., and Gill, B.S. (2022). Deciphering the Mechanism of Glyphosate Resistance in Amaranthus palmeri by Cytogenomics. Cytogenet. Genome Res. *161*, 578–584. 10.1159/000521409.
55. Koo, D.H., Sathishraj, R., Nakka, S., Ju, Y., Nandula, V.K., Jugulam, M., Friebe, B., and Gill, B.S. (2023). Extrachromosomal circular DNA–mediated spread of herbicide resistance in interspecific hybrids of pigweed. Plant Physiol. *193*, 229–233. 10.1093/plphys/kiad281.
56. Koo, D.H., Ju, Y., Putta, K., Sathishraj, R., Burgos, N., Jugulam, M., Friebe, B., and Gill, B.S. (2023). Extrachromosomal DNA-mediated glyphosate resistance in Italian ryegrass. Pest Manag. Sci., 25–29. 10.1002/ps.7626.
57. Zhang, C. (2023). Subtelomeric 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) copy number variation confers glyphosate resistance in Eleusine indica. Res. Sq., 0–24. 10.1038/s41467-023-40407-6.
58. Jugulam, M., Niehues, K., Godar, A.S., Koo, D.H., Danilova, T., Friebe, B., Sehgal, S., Varanasi, V.K., Wiersma, A., Westra, P., et al. (2014). Tandem amplification of a chromosomal segment harboring 5-Enolpyruvylshikimate-3-Phosphate synthase locus confers glyphosate resistance in Kochia scoparia. Plant Physiol. *166*, 1200–1207. 10.1104/pp.114.242826.
59. Singh, V., Etheredge, L., McGinty, J., Morgan, G., and Bagavathiannan, M. (2020). First case of glyphosate resistance in weedy sunflower (Helianthus annuus). Pest Manag. Sci. *76*, 3685–3692. 10.1002/ps.5917.
60. Adu-Yeboah, P., Malone, J.M., Fleet, B., Gill, G., and Preston, C. (2020). EPSPS gene amplification confers resistance to glyphosate

resistant populations of Hordeum glaucum Stued (northern barley grass) in South Australia. Pest Manag. Sci. *76*, 1214–1221. 10.1002/ps.5671.

61. Ngo, T.D., Malone, J.M., Boutsalis, P., Gill, G., and Preston, C. (2018). EPSPS gene amplification conferring resistance to glyphosate in windmill grass (Chloris truncata) in Australia. Pest Manag. Sci. *74*, 1101–1108. 10.1002/ps.4573.
62. Patterson, E.L., Pettinga, D.J., Ravet, K., Neve, P., and Gaines, T.A. (2018). Glyphosate resistance and EPSPS gene duplication: Convergent evolution in multiple plant species. J. Hered. *109*, 117–125. 10.1093/jhered/esx087.
63. Shyam, C., Borgato, E.A., Peterson, D.E., Dille, J.A., and Jugulam, M. (2021). Predominance of Metabolic Resistance in a Six-Way-Resistant Palmer Amaranth (Amaranthus palmeri) Population. Front. Plant Sci. *11*, 1–12. 10.3389/fpls.2020.614618.
64. Kaya-Altop, E., Jabran, K., Pala, F., and Mennan, H. (2024). Multiple resistance to EPSPS and ALS inhibitors in Palmer amaranth (Amaranthus palmeri) identified in Turkey. Weed Res. 10.1111/wre.12618.
65. 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., et al. (2010). Gene amplification confers glyphosate resistance in Amaranthus palmeri. Proc. Natl. Acad. Sci. U. S. A. *107*, 1029–1034. 10.1073/pnas.0906649107.
66. Pagnoncelli Jr., F.D.B., Losada, F.B., Alvear, M.J.G., Gonzalez-Andujar, J.L., Trezzi, M.M., Bittencourt, H.V.H., and Salomão, H.M. (2024). Response characterization and target site mechanism study in glyphosate-resistant populations of Lolium multiflorum L. from Brazil. Pestic. Biochem. Physiol. *198*, 105737. https://doi.org/10.1016/j.pestbp.2023.105737.
67. Barua, R., Malone, J., Boutsalis, P., Gill, G., and Preston, C. (2022). Inheritance and mechanism of glyphosate resistance in annual bluegrass (Poa annua L.). Pest Manag. Sci. *78*, 1377–1385. 10.1002/ps.6754.
68. Shergill, L.S., Bish, M.D., Jugulam, M., and Bradley, K.W. (2018). Molecular and Physiological Characterization of six-way resistance in Amaranthus tuberculatus var.rudis biotype from Missouri. Pest Manag. Sci. *74*, 2688–2698.
69. Yanniccari, M., Gómez-Lobato, M.E., Istilart, C., Natalucci, C., Giménez, D.O., and Castro, A.M. (2017). Mechanism of resistance to glyphosate in Lolium perenne from Argentina. Front. Ecol. Evol. *5*, 1–8. 10.3389/fevo.2017.00123.
70. Zhang, C., Feng, L., He, T.T., Yang, C.H., Chen, G.Q., and Tian, X.S. (2015). Investigating the mechanisms of glyphosate resistance in goosegrass (Eleusine indica) population from South China. J. Integr. Agric. *14*, 909–918. 10.1016/S2095-3119(14)60890-X.
71. Mei, Y., Xu, Y., Wang, S., Qiu, L., and Zheng, M. (2018). Investigation of glyphosate resistance levels and target-site based resistance (TSR) mechanisms in Conyza canadensis (L.) from apple orchards around areas of Bohai seas and Loess Plateau in China. Pestic. Biochem. Physiol. *146*, 7–12. 10.1016/j.pestbp.2017.12.008.
72. Vila-Aiub, M., Balbi, M.C., Distéfano, A.J., Fernández, L., Hopp, E., Yu, Q., and Powles, S.B. (2011). Glyphosate resistance in perennial Sorghum halepense (Johnsongrass), endowed by reduced glyphosate translocation and leaf uptake. Pest Manag. Sci. *68*, 430–436. 10.1002/ps.2286.
73. Michitte, P., De Prado, R., Espinoza, N., Pedro Ruiz-Santaella, J., and Gauvrit, C. (2007). Mechanisms of Resistance to Glyphosate in a Ryegrass (Lolium Multiflorum) Biotype from Chile . Weed Sci. *55*, 435–440. 10.1614/ws-06-167.1.
74. Palma-Bautista, C., Vázquez-Garciá, J.G., Domínguez-Valenzuela, J.A., Ferreira Mendes, K., Alcántara De La Cruz, R., Torra, J., and

De Prado, R. (2021). Non-Target-Site Resistance Mechanisms Endow Multiple Herbicide Resistance to Five Mechanisms of Action in Conyza bonariensis. J. Agric. Food Chem. *69*, 14792–14801. 10.1021/acs.jafc.1c04279.

75. Peng, Y., Abercrombie, L.L.G., Yuan, J.S., Riggins, C.W., Sammons, R.D., Tranel, P.J., and Stewart, C.N. (2010). Characterization of the horseweed (Conyza canadensis) transcriptome using GS-FLX 454 pyrosequencing and its application for expression analysis of candidate non-target herbicide resistance genes. Pest Manag. Sci. *66*, 1053–1062. 10.1002/ps.2004.
76. Tani, E., Chachalis, D., and Travlos, I.S. (2015). A Glyphosate Resistance Mechanism in Conyza canadensis Involves Synchronization of EPSPS and ABC-transporter Genes. Plant Mol. Biol. Report. *33*, 1721–1730. 10.1007/s11105-015-0868-8.
77. Moretti, M.L., Alárcon-Reverte, R., Pearce, S., Morran, S., and Hanson, B.D. (2017). Transcription of putative tonoplast transporters in response to glyphosate and paraquat stress in Conyza bonariensis and Conyza canadensis and selection of reference genes for qRT-PCR. PLoS One *12*, 1–16. 10.1371/journal.pone.0180794.
78. Schneider, T., Rizzardi, M.A., Grando, M.F., Ceccon, C.C., Nunes, A.L., and Didone, D. (2021). Expression of genes related to the mechanism of resistance of conyza sumatrensis to glyphosate. Acta Sci. - Agron. *43*, 1–8. 10.4025/actasciagron.v43i1.44824.
79. Gerakari, M., Cheimona, N., Tani, E., Travlos, I., Chachalis, D., Loddo, D., Mathiassen, S.K., Gitsopoulos, T.K., Scarabel, L., Panozzo, S., et al. (2022). Biochemical and rapid molecular analyses to identify glyphosate resistance in lolium spp. Agronomy *12*, 1–12. 10.3390/agronomy12010040.
80. Pan, L., Yu, Q., Wang, J., Han, H., Mao, L., Nyporko, A., Maguza, A., Fan, L., Bai, L., and Powles, S. (2021). An ABCC-type transporter endowing glyphosate resistance in plants. Proc. Natl. Acad. Sci. U. S. A. *118*, 1–11. 10.1073/pnas.2100136118.
81. Vazquez-Garcia, J.G., Palma-Bautista, C., Rojano-Delgado, A.M., De Prado, R., and Menendez, J. (2020). The first case of glyphosate resistance in johnsongrass (Sorghum halepense (l.) pers.) in Europe. Plants *9*, 1–11. 10.3390/plants9030313.
82. Pan, L., Yu, Q., Han, H., Mao, L.L., Nyporko, A., Fan, L., Bai, L., and Powles, S.B. (2019). Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in Echinochloa colona. Plant Physiol. *181*, pp.00979.2019. 10.1104/pp.19.00979.
83. Zhou, F.Y., Han, H., Han, Y.J., Nyporko, A., Yu, Q., Beckie, H.J., and Powles, S.B. (2023). Aldo-keto reductase may contribute to glyphosate resistance in Lolium rigidum. Pest Manag. Sci. *79*, 1528–1537. 10.1002/ps.7325.
84. Leslie, T., and 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, Genet. *4*, 2035–2047. 10.1534/g3.114.013508.
85. Van Etten, M., Lee, K.M., Chang, S.M., and Baucom, R.S. (2020). Parallel and nonparallel genomic responses contribute to herbicide resistance in Ipomoea purpurea, a common agricultural weed 10.1371/journal.pgen.1008593.
86. Laforest, M., Martin, S.L., Bisaillon, K., Soufiane, B., Meloche, S., and Page, E. (2020). A chromosome-scale draft sequence of the Canada fleabane genome. Pest Manag. Sci. *76*, 2158–2169. 10.1002/ps.5753.
87. Piasecki, C., Yang, Y., Benemann, D.P., Kremer, F.S., Galli, V., Millwood, R.J., Cechin, J., Agostinetto, D., Maia, L.C., Vargas, L., et al. (2019). Transcriptomic analysis identifies new non-target site glyphosate-resistance genes in conyza bonariensis. Plants *8*. 10.3390/plants8060157.

88. Piasecki, C., Carvalho, I.R., Cechin, J., Goulart, F.A.P., da Maia, L.C., Agostinetto, D., Caverzan, A., Stewart, C.N., and Vargas, L. (2019). Oxidative stress and differential antioxidant enzyme activity in glyphosate-resistant and-sensitive hairy fleabane in response to glyphosate treatment. Bragantia *78*, 379–396. 10.1590/1678-4499.20180289.
89. Harre, N.T., Nie, H., Jiang, Y., and Young, B.G. (2018). Differential antioxidant enzyme activity in rapid-response glyphosate-resistant Ambrosia trifida. Pest Manag. Sci. *74*, 2125–2132. 10.1002/ps.4909.

Appendix 2: *C.reinhardtii* DNA extractions protocol

Materials& Solutions:

Genomic extraction buffer (GE)

PVP-40

Sodium bisulfite (SB)

Glass beads

KAc

TE Buffer

RNase (working stock @ 2mg/ml)

Isopropanol

70%Ethanol

5ml,1ml, 200 µl **sterile** tips

15ml and 1.5 ml **sterile** tubes.

Before getting started:

- Keep Isopropanol at -20°C
- Start water bath at 65°C
- Make up GE buffer: 7.5 mg PVP 40 & 3.6 mg of SB/ ml of GE
- Make up TER buffer: 25 ul of RNase/ ml TE buffer

Protocol:

1. Cell lysing step	a. Snap freeze in liquid N (2 mins) and thaw (2 mins). Repeat X3 b. add glass beads and 1ml of GE buffer Freshly made up in each sample c. bead beater max speed for 20 secs d. Incubate at 65°C for 1hour. Spin down and turn temperature of water bath on 50°C
2. Isolating *Nucleic acids* separating debris in the pellet from the *Nucleic acids in the supernatant*	a. add 333 ml of KAc solution b. Vortex for 2 mins c. Centrifuge at 10,000 rcf for 15 mins
3. Precipitation of *Nucleic acids*	a. Aliquot 550 µl of ice cold Isopropanol into 1.5 ml tubes b. Add 1ml of supernatant c. Mix by inversion x6 d. Incubate 10 min at room temperature
4. Pelleting *Nucleic acids*	d. *Centrifuge at 18,000 rcf for 15 mins*
5. Washing the Pellet Using Ethanol to remove contaminants from pelleted nucleic acids	a. Discard supernatant b. Add 500 µl of 70% ethanol c. Mix by inversion d. Centrifuge at 2,5000 rcf for 5 mins.
6. Dry the Pellet	a. Discard supernatant

Removing traces of Ethanol and contaminant without over drying the pellets.	b. Remove remaining Ethanol using a pipette 100 µl
7. Eluting *Nucleic acids*	• add 75 µl of TER to an ethanol free pellet
8. Digesting RNA	a) Incubate at 50°C for 1hour. Mix every 15 mins b) Centrifuge for 5 mins at 16,000 rcf to pellet polysaccharides. Transfer supernatant in final storage tube
9. Storage conditions	• Store at 4°C for immediate use or at -20°C for longer term

Appendix 3: List of variants susceptible to support glyphosate resistance from the SR experiment (Chapter 3)

List of variants retained after filtering for high quality variants susceptible to support glyphosate resistance. For variant calling pipeline and filters details refer to Chapter 2

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
05	3,685,946	AGGGGAAGGGGAGGGGGAGGCGGGT	AGGGGGAAGGGGAGGGGGAGGCGGGT	ins	425.0	GR_A	156	1.00
						GS_A		
						GR_B	156	1.00
						GS_B		
						GR_C	140	1.00
						GS_C		
01	19,761	C	T	snp	261.0	GR_B	160	1.00
						GS_B	160	0.15
						GR_C	47	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_C	160	0.22
	1,555,693	G			1,858.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.04
	2,390,845	C			5,301.0	GR_B	138	1.00
						GS_B	138	0.14
						GR_C	138	0.97
						GS_C	138	0.12
	3,980,468	TCA	TA	del	8,005.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	4,097,932	GACACACACACACACACACACACACACACACACAT	GACACACACACACACACACACACACACACACACACAT	ins	2,234.0	GR_B	143	1.00
						GS_B	143	0.00
						GR_C	143	1.00
						GS_C	143	0.00
	4,740,525	CCAGCAGCAGCAGCAGCAGCAGCAGC	CCAGCAGCAGCAGCAGCAGCAGCAGCAGC		4,434.0	GR_B	160	0.96
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	5,393,698	C	A	snp	7,890.0	GR_B	147	1.00
						GS_B	147	0.00
						GR_C	147	1.00
						GS_C	147	0.00
	6,475,904			ins	1,288.0	GR_B	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		CGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGCCGGGGCCGGGCCGGGGCCGGGCA	CGGCCGGGGCCGGGGCCGGGGCCGGGCCGGGGCCGGGCCGGGGCCGGGCCGGGGCCGGGCA			GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	7,148,825	GGCACACG	GG	del	5,360.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	7,304,996	TG	TTCGTTAA	complex	7,658.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
02	420,383	CGTGTGTGTGTGTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGC	ins	1,231.0	GR_B	145	1.00
						GS_B	145	0.00
						GR_C	145	1.00
						GS_C	145	0.00
	1,019,958	G	A	snp	7,583.0	GR_B	149	1.00
						GS_B	149	0.00
						GR_C	149	1.00
						GS_C	149	0.00
	3,116,800	T	G		978.0	GR_B	142	1.00
						GS_B	142	0.00
						GR_C	142	1.00
						GS_C	142	0.00
	3,567,631			ins	5,036.0	GR_B	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		TCACACACACATCACACACACA CACACACACA CACAT	TCACACACACA CACACACACA CACAT			GS_B	160	0.00
						GR_C	160	0.98
						GS_C	160	0.00
	5,354,377	GCACACACAC ACACACACAC G	GCACACACAC ACACACACAC ACG		2,957.0	GR_B	137	1.00
						GS_B	137	0.00
						GR_C	137	0.96
						GS_C	112	0.49
	5,938,736	ATGTGTGTGTG TGTGTGTGTGT GTGTGC	ATGTGTGTGTG TGTGTGTGTGT GTGTGTGC		3,245.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	6,585,231		CGTGTGTGTGT GTGTGTGTGTG		1,691.0	GR_B	139	1.00
						GS_B	139	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		CGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	TGTGTGTGTGTGTTA			GR_C	139	1.00
						GS_C	139	0.00
03	4,890,271	CGGCC	CGCC	del	7,630.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	4,939,424	TCACCACCACCACCACCACCACCACCACCACACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCTT	TCACCACCACCACCACCACCACCACCACCACACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCTT	ins	691.0	GR_B	139	1.00
						GS_B	139	0.00
						GR_C	139	1.00
						GS_C	139	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
04	32	A	C	snp	939.0	GR_B	143	0.90
						GS_B	135	0.25
						GR_C	143	0.96
						GS_C	143	0.20
	82,722	TGCAGCAGCAGCAGCAGCAG CAG	TGCAGCAGCAGCAGCAGCAG CAGCAG	ins	4,243.0	GR_B	160	0.89
						GS_B	160	0.00
						GR_C	160	0.95
						GS_C	160	0.00
	2,493,939	GCCCCCCCCCC TCTTA	GCCCCCCCCCC CTCTTA		4,585.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	3,658,937			del	283.0	GR_B	140	0.92

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		GCACACACACACACACACACACACACACACACAG	GCACACACACACACACACACACACACACACAG			GS_B	140	0.00
						GR_C	117	1.00
						GS_C	98	0.00
	3,942,575	AACACACACACACACACACACACACACAT	AACACACACACACACACACACACACACAT		2,037.0	GR_B	160	0.97
						GS_B	160	0.00
						GR_C	160	0.98
						GS_C	160	0.00
05	635,242	TCCCCCCCTGCG	TCCCCCCCCTGCG	ins	2,952.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
06	384,677	CGTGTGTGTGTGTGTGTGTGTG	CGTGTGTGTGTGTGTGTGTG		2,784.0	GR_B	149	0.98
						GS_B	149	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		TGTGTGTGTGTGTGTA	TGTGTGTGTGTGTGTA			GR_C	149	0.96
						GS_C	149	0.00
	1,669,850	GGTGTGTGTGTGGTGTGGGGGGGG	GGTGTGTGTGTGGTGTGGGGGGGG		498.0	GR_B	148	1.00
						GS_B	1	0.00
						GR_C	148	1.00
						GS_C	2	0.00
	2,083,864	GCCACCACCACCACCACCACCACCACCT	GCCACCACCACCACCACCACCT	del	2,534.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	3,385,106	GGTGTGTGTGTGTGTGTGTGTGA	GGTGTGTGTGTGTGTGTGTGA	ins	1,837.0	GR_B	160	0.97
						GS_B	160	0.00
						GR_C	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_C	160	0.00
	6,463,824	CTGTGTGTGTGTGTGTGTGC	CTGTGTGTGTGTGTGTGTGC		4,035.0	GR_B	160	0.98
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	7,000,602	GGCTGCTGCTGCTGCTGCTGCTGCTG	GGCTGCTGCTGCTGCTGCTGCTGCTG		4,282.0	GR_B	155	1.00
						GS_B	155	0.00
						GR_C	155	0.98
						GS_C	155	0.00
	8,266,021	CCACACACACACACACACACACACACG	CCACACACACACACACACACG	del	4,929.0	GR_B	160	1.00
						GS_B	160	0.03
						GR_C	160	1.00
						GS_C	160	0.03

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
07	914,701	ATGTGTGTGTGTGTGTGTTTA	ATGTGTGTGTGTGTGTGTTTA	ins	1,724.0	GR_B	144	1.00
						GS_B	144	0.00
						GR_C	144	1.00
						GS_C	144	0.00
	1,010,355	CTCGTCGTCGTCGTCGTCGTCGTCGTCGG	CTCGTCGTCGTCGTCGTCGTCGTCGTCGG		4,395.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	1,781,173	CGCGGCTGTAGCGGCTGTAGCGGCTGTAG	CGCGGCTGTAGCGGCTGTAGCGGCTGTAGCGGCTGTAG		4,312.0	GR_B	160	0.96
						GS_B	160	0.00
						GR_C	160	0.88
						GS_C	160	0.00
	3,860,633			del	713.0	GR_B	145	0.94

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		ATGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGC	ATGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGC			GS_B	145	0.00
						GR_C	145	1.00
						GS_C	145	0.10
	4,678,848	AACACACACA CACACACACA CACACACACA CACACACAAC ACACG	AACACACACA CACACACACA CACACACACA CACACACACA ACACACG	ins	624.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.94
						GS_C	160	0.00
08	340,984	C	G	snp	5,954.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	415,616	A	C		5,087.0	GR_B	143	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_B	143	0.00
						GR_C	143	1.00
						GS_C	143	0.01
	3,215,241	GCACACACACACACACACACACACACACACT	GCACACACACACACACACACACACACACACT	ins	1,678.0	GR_B	139	0.91
						GS_B	139	0.00
						GR_C	139	1.00
						GS_C	139	0.00
	3,918,420	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG		4,750.0	GR_B	160	0.97
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	4,439,267	A	C	snp	4,891.0	GR_B	143	1.00
						GS_B	143	0.01

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GR_C	143	1.00
						GS_C	143	0.00
09	810,582	C	G	ins	7,979.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	1,439,622	GCACACACACACACACACACAT	GCACACACACACACACACAT		891.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	2,467,744	ATGTGTGTGTGTGTGTGTGTGTGA	ATGTGTGTGTGTGTGTGTGTGA		3,923.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.98

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_C	160	0.03
	5,638,715	TCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCG	TCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCG		533.0	GR_B	148	1.00
						GS_B	148	0.00
						GR_C	148	1.00
						GS_C	148	0.00
	6,380,673	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA		1,148.0	GR_B	160	0.96
						GS_B	160	0.00
						GR_C	160	0.92
						GS_C	160	0.00
10	900,763	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTTG	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG		1,652.0	GR_B	142	0.94
						GS_B	142	0.00
						GR_C	142	0.97
						GS_C	142	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	2,116,960	ACCCCCCCCCCCCACCGCCGCCCCTCCCCAC	ACCCCCCCCCCCCCCCACCGCCGCCCTCCCCAC		381.0	GR_B	137	1.00
						GS_B		
						GR_C	160	0.92
						GS_C		
	4,436,009	G	T	snp	5,768.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.99
						GS_C	160	0.00
	4,528,327	TGTGCGTGCGTGCGTGCGTGTGCGTGCGTGTGCGTGCGTGTGCGTGCG	TGTGCGTGCGTGCGTGCGTGTGCGTGCGTGTGCGTGCGTGTGCGTGCG	del	522.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		GCGTGCGTGCG						
	6,215,827	GCCCCCCCCCCACACACACACACACACCCT	GCCCCCCCCCCCCACACACACACACACCCT	ins	1,729.0	GR_B	160	0.94
						GS_B	160	0.00
						GR_C	160	0.95
						GS_C	160	0.00
	6,518,000	T	G	snp	2,451.0	GR_B	143	1.00
						GS_B	143	0.00
						GR_C	143	0.98
						GS_C	143	0.11
	6,719,219	TTTTAGGGTTTAGGGT	TTTTAGGGTTTTAGGGT	ins	2,110.0	GR_B	160	0.98
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.01

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
11	1,039,723	GACACACACACACACACACACACACACACACACACAG	GACACACACACACACACACACACACACACACACACAG		2,925.0	GR_B	141	0.97
						GS_B	141	0.02
						GR_C	141	0.97
						GS_C	141	0.00
	3,189,268	AGGCCGCGGCCGGGGCCGCGGCCGGGGCCGT	AGGCCGCGGCCGGGGCCGCGGCCGGGGCCGT		2,974.0	GR_B	146	1.00
						GS_B	146	0.00
						GR_C	146	1.00
						GS_C	146	0.00
12	1,296,987	T	C	snp	6,575.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00
	1,535,558			complex	1,513.0	GR_B	144	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		GTTTATCAACACG	GCTGTGGGCACGGGTGGGAGCCACACACATCGGGACTACGCTTTCCTGCCCCTTGCCATGTTCCATCACGAACACG			GS_B	144	0.00
						GR_C	144	1.00
						GS_C	144	0.00
	7,880,011	CTTGGGTGG	CAGGTACTGTCGGT		7,088.0	GR_B	149	1.00
						GS_B	149	0.00
						GR_C	149	1.00
						GS_C	149	0.00
13	1,266,357	GCACACACACACACACACACACACATAC	GCACACACACACACACACACACACACATAC	ins	1,033.0	GR_B	160	0.94
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	3,411,519	TGGGGGGGGGGAAC	TGGGGGGGGGGGAAC		3,093.0	GR_B	160	0.96
						GS_B	160	0.00
						GR_C	160	0.95
						GS_C	160	0.00
	4,149,461	A	C	snp	6,623.0	GR_B	154	1.00
						GS_B	154	0.00
						GR_C	154	1.00
						GS_C	154	0.00
	4,798,082	CAAGG	CG	del	6,316.0	GR_B	147	1.00
						GS_B	147	0.00
						GR_C	147	1.00
						GS_C	147	0.00
14	422,685			ins	1,301.0	GR_B	160	0.94

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		CCACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACG	CCACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACACG			GS_B	160	0.00
						GR_C	160	0.94
						GS_C	160	0.00
	1,532,630	CGGGGGGGGG GGCGG	CGGGGGGGGG GGGCGG		1,680.0	GR_B	160	0.92
						GS_B	160	0.00
						GR_C	160	0.92
						GS_C	160	0.00
	2,276,825	CGGGGGGGGG GGCGCTC	CGGGGGGGGG GGGCGCTC		548.0	GR_B	139	0.88
						GS_B	0	
						GR_C	139	1.00
						GS_C	0	
	3,850,348			del	946.0	GR_B	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		CTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGGC	CTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGGC			GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.04
	4,186,933	G	A		531.0	GR_B	0	0.94
						GS_B	160	0.18
						GR_C	160	0.99
						GS_C	160	0.34
	4,188,448	A	G	snp	0.0	GR_B	0	1.00
						GS_B	160	0.06
						GR_C	0	1.00
						GS_C	160	0.15
	4,188,777	T			220.0	GR_B	0	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_B	160	0.13
						GR_C	160	1.00
						GS_C	160	0.23
15	3,555,493	ATATTATTC	TTATTATTC		1,988.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	0.97
						GS_C	160	0.00
16	512,172	AACACACACACACACACACG	AACACACACACACACACACG	ins	1,913.0	GR_B	160	0.96
						GS_B	160	0.00
						GR_C	160	1.00
						GS_C	160	0.02
	2,260,411	GGCTGCTGCTGCTGCTGCTG			5,379.0	GR_B	148	0.88
						GS_B	148	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
			GGCTGCTGCTGCTGCTGCTGC TG			GR_C	148	0.90
						GS_C	148	0.00
	2,815,045	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGC	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGC		3,114.0	GR_B	160	0.98
						GS_B	160	0.00
						GR_C	160	0.97
						GS_C	160	0.02
	4,948,956	CCACACACACACACACACACACACACACACACACACAT	CCACACACACACACACACACACACACACACACACACAT	del	2,254.0	GR_B	160	1.00
						GS_B	160	0.06
						GR_C	160	1.00
						GS_C	160	0.02
	5,230,480	GTA	GTCTTGCATTCGCATTCACTCGCAAGCACTTCGTTA	ins	3,158.0	GR_B	160	1.00
						GS_B	160	0.00
						GR_C	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_C	160	0.00
	7,716,527	AGGGGGGGGGCGTTGT	AGGGGGGGGGGCGTTGT		895.0	GR_B	140	1.00
						GS_B	140	0.02
						GR_C	140	0.93
						GS_C	140	0.00
17	2,372,299	A	G	snp	7,478.0	GR_B	148	1.00
						GS_B	148	0.00
						GR_C	148	1.00
						GS_C	148	0.00
	4,182,600	ATGTGTGTGTGTGTGTGTGTGTA	ATGTGTGTGTGTGTGTGTA	del	4,254.0	GR_B	145	0.98
						GS_B	145	0.03
						GR_C	145	1.00
						GS_C	145	0.01

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	5,020,066	CTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT	CTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT		2,910.0	GR_B	143	1.00
						GS_B	143	0.00
						GR_C	143	1.00
						GS_C	143	0.00
	6,471,328	TCAGCAGCAGCAGCAGG	TCATCAGCAGCATCATCAGCAGCATCATCAGCAGGCCGGCGGCAAGGGCGGCAAGCAGCAGCAGCATCAGCAGCAGG	complex	67.0	GR_B	67	1.00
						GS_B	141	0.00
						GR_C	0	1.00
						GS_C	141	0.00
01	564,397	GACACACACACACACACACACACACACACACACACACACACACT	GACACACACACACACACACACACACACACACACACACACACACACT	ins	758.0	GR_A	148	0.97
						GS_A	148	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	2,397,527	CACACAACACAACACAACACAACACAACACAACACAACACAACACAACACAACAAACACACAG	CACACAACACAACACAACACAACACAACACAACACAACACAACACAACACAACACAACACAACAAACACACAG		1,440.0	GR_A	141	1.00
						GS_A	141	0.00
	3,152,613	ACCCCCCCCCTGG	ACCCCCCCCCCTGG		1,989.0	GR_A	160	0.97
						GS_A	160	0.00
	3,252,015	ATAACCCCACGATGT	AT	del	204.0	GR_A	142	1.00
						GS_A	142	0.00
	4,839,802	T	G	snp	317.0	GR_A	147	1.00
						GS_A	147	0.03
	5,621,006	AGTGTGTGTGTGTGTGTGTGTGTGTA	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	ins	283.0	GR_A	147	1.00
						GS_A	147	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
02	423,838	T	G	snp	2,759.0	GR_C	160	1.00
						GS_C	160	0.01
	994,394	TGGGGGGGGGGGGCAGAGGCA	TGGGGGGGGGGGGCAGAGGCA	ins	390.0	GR_A	160	1.00
						GS_A	160	0.00
	1,053,941	GCACACACACACACACACACACAT	GCACACACACACACACACACACAT		2,117.0	GR_C	160	0.99
						GS_C	160	0.01
	1,835,505	AACACACACACACACACACACACG	AACACACACACACACACACACACG		2,260.0	GR_A	160	1.00
						GS_A	160	0.00
	5,399,695	ACCACGCCACGCCACGCCACGCCACGCCACGCCACGCCACGCCAC	ACCACGCCACGCCACGCCACGCCACGCCACGCCACGCCAC	del	1,552.0	GR_A	143	1.00
						GS_A	143	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		GCCACGCCACGC	GCCACGCCACGCCACGC					
	6,249,627	TTGTGTGTGTGTGTGTGTGA	TTGTGTGTGTGTGTGTGTGTGA		2,147.0	GR_A	144	1.00
						GS_A	144	0.00
	6,857,626	TAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	TAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA	ins	1,728.0	GR_A	141	0.92
						GS_A	141	0.00
	8,514,074	CGTGTGTGTGTGTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGTGC		820.0	GR_A	145	0.97
						GS_A	145	0.00
03	1,385,028				1,579.0	GR_A	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		TACACACACACACACACACACG	TACACACACACACACACACG			GS_A	160	0.00
	6,245,369	CCAGCAGCAGCAGCAGCAGCAGCAGCAGA	CCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGA		2,943.0	GR_A	160	1.00
						GS_A	160	0.00
	8,524,378	GCACACACACACACACACACACACACACACACACG	GCACACACACACACACACACACACACACACACG		397.0	GR_A	160	0.91
						GS_A	160	0.00
04	189,699	ACAGCAGCAGCAGCAGCAGCAGG	ACAGCAGCAGCAGCAGCAGCAGG		2,215.0	GR_A	160	1.00
						GS_A	160	0.00
	286,847				850.0	GR_A	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		AACACACACACACACACACACACACACACG	AACACACACACACACACACACACACACG			GS_A	160	0.00
	286,887	TACACACACACTACACACACACACACACACACACT	TACACACACACACACACACACACACT	del	1,310.0	GR_A	160	1.00
						GS_A	160	0.02
	839,723	C	T	snp	2,745.0	GR_B	140	0.93
						GS_B	140	0.00
	1,267,796				3,565.0	GR_A	154	1.00
						GS_A	154	0.00
	1,397,506	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC	ins	2,672.0	GR_A	143	0.98
						GS_A	143	0.00
	1,863,043	G	C	snp	1,685.0	GR_A	146	1.00
						GS_A	146	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	2,087,427	CTGTGTGTGTGTGTGTGTGTGTGGGTGTGGGTGC	CTGTGTGTGTGTGTGTGTGTGTGTGGGTGTGGGTGC	ins	280.0	GR_C	155	1.00
						GS_C	155	0.00
	2,172,588	TGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG	TGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG	del	475.0	GR_B	160	1.00
						GS_B	160	0.00
	3,428,879	AACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAG	AACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAG		312.0	GR_A	141	1.00
						GS_A	141	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
05	48,804	CCA	CCCCCACA	ins	958.0	GR_A	160	0.91
						GS_A	160	0.00
	433,737	AGGGGGGGGG TGATT	AGGGGGGGGG GTGATT		232.0	GR_A	149	0.92
						GS_A	149	0.00
	685,392	GCCCCCCCCC CCCGCCTAA	GCCCCCCCCC CCGCCTAA	del	2,087.0	GR_B	160	0.99
						GS_B	160	0.06
	922,926	CGTTGTTGTTG TTGTTGTTGTT GTTGTTGTTGT TGTTGTTGTTG TTGTTGTTGTT GTTGTTGTTGT TGTTGTTGTTG TTGTTGTTGTT GTTGTTGCTG	CGTTGTTGTTG TTGTTGTTGTT GTTGTTGTTGT TGTTGTTGTTG TTGTTGTTGTT GTTGTTGTTGT TGTTGTTGTTG TTGTTGTTGTT GTTGTTGTTGC TG	ins	1,298.0	GR_A	160	1.00
						GS_A	160	0.03

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	1,048,297	TACACACACACTACACACACACACACACACACACACACACACAG	TACACACACACACACACACACACACACACACACACACACACACAG		613.0	GR_A	142	0.93
						GS_A	142	0.00
	1,129,258	TACACACACACTACACACT	TACACACACACACT	del	1,766.0	GR_A	160	1.00
						GS_A	160	0.00
	1,654,763	TCGCCC	TCGCCCGCCC	ins	1,497.0	GR_A	160	1.00
						GS_A	160	0.00
	2,880,401	C	T	snp	3,582.0	GR_A	160	0.99
						GS_A	160	0.00
06	1,413,415	GGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	GGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	ins	2,104.0	GR_C	160	0.91
						GS_C	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	1,471,155	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA		1,527.0	GR_C	143	0.95
						GS_C	143	0.00
	2,084,732	AACACACACACACACACACAT	AACACACACACACACACACAT		1,214.0	GR_C	147	1.00
						GS_C	147	0.00
	3,699,555	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG		2,769.0	GR_A	160	0.94
						GS_A	160	0.00
	4,152,923	ACCCCCCCCCCCCGGCTCA	ACCCCCCCCCCCGGCTCA		2,265.0	GR_A	143	1.00
						GS_A	143	0.03
	4,726,506	CATATATATATATATATATG	CATATATATATATATATG	del	4,177.0	GR_B	160	1.00
						GS_B	160	0.02
	5,210,759	C	G	snp	4,572.0	GR_A	160	1.00
						GS_A	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
07	6,441,214	TGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAC	TGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAC	ins	2,540.0	GR_A	145	0.99
						GS_A	145	0.00
	6,704,670	AACACACACACACACACACACACACACACACACACC	AACACACACACACACACACACACACACACACACACAACC		1,256.0	GR_A	160	0.98
						GS_A	160	0.00
	7,342,788	AGGGGGGGGGGTAGGAA	AGGGGGGGGGTAGGAA	del	1,784.0	GR_A	160	1.00
						GS_A	160	0.02
	8,621,457	A	C	snp	4,302.0	GR_A	151	1.00
						GS_A	151	0.00
	431,386	C	G		3,094.0	GR_A	151	1.00
						GS_A	151	0.00
	3,582,932			del	1,362.0	GR_A	148	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		GCACACACACACACACACACACACACACACACACACACACACACACCT	GCACACACACACACACACACACACACACACACACACACCT			GS_A	148	0.00
	3,618,668	AGGGGGGGGGGGTTGTA	AGGGGGGGGGGGGTTGTA	ins	381.0	GR_A	139	0.94
						GS_A	139	0.00
	4,133,815	C	T	snp	3,208.0	GR_A	157	1.00
						GS_A	157	0.01
	4,166,695	TACACACACACTACACAT	TACACACACACACACAT	ins	2,332.0	GR_A	160	0.98
						GS_A	160	0.00
	4,729,476	GCACACACACACACACACACACACACACACAT	GCACACACACACACACACACACACACACACAT		2,374.0	GR_A	160	0.99
						GS_A	160	0.00
	4,769,875	G	A	snp	1,770.0	GR_A	149	1.00
						GS_A	149	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
08	5,053,746	GCACACACAC ACACACACAC ACACACAG	GCACACACAC ACACACACAC ACACACACAG	ins	399.0	GR_A	146	0.95
						GS_A	146	0.00
	6,126,432	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTTA	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTTA		223.0	GR_A	160	1.00
						GS_A	160	0.00
	286,819	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACG	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACG		227.0	GR_C	160	0.92
						GS_C	0	
	767,586	CGGGGGGGGG GCGCTGT	CGGGGGGGGG GGCGCTGT		1,687.0	GR_A	144	0.98
						GS_A	144	0.01
	1,305,433				775.0	GR_A	160	0.90

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA			GS_A	160	0.00
	1,516,859	T	G	snp	1,977.0	GR_B	160	1.00
						GS_B	160	0.02
	1,621,824	ACCCCCCCCCCATCCCCCCCAT	ACCCCCCCCCCCATCCCCCCCAT		719.0	GR_A	142	0.96
						GS_A	142	0.00
	1,774,937	CCCGCCACCGCCACCGCCACCGCCACCGCCAC	CCCGCCACCGCCACCGCCACCGCCACCGCCAC	ins	1,904.0	GR_A	160	0.92
						GS_A	160	0.00
	3,012,012	ACCCCCCCCCAGG	ACCCCCCCCCCAGG		370.0	GR_A	143	1.00
						GS_A	143	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
09	1,183,240	TCACACACACACACACACACACACACACACACACACAT	TCACACACACACACACACACACACACACACACAT	del	1,767.0	GR_A	150	1.00
						GS_A	150	0.02
	3,186,690	GTATGTGTGTGTGTGTGTGTGTGTTTGG	GGTATGTGTGTGTGTGTGTGTGTTTGG	complex	263.0	GR_A	23	1.00
						GS_A	139	0.00
	3,621,877	C	G	snp	209.0	GR_B	119	0.89
						GS_B	133	0.30
	3,801,124	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	ins	3,213.0	GR_A	142	0.94
						GS_A	142	0.00
	4,096,128	CGTGTGTGTGTGTGTGTGTGTGTGTTG	CGTGTGTGTGTGTGTGTGTGTGTGTTG		2,113.0	GR_A	160	0.96
						GS_A	160	0.00
	4,302,169	G	A	snp	93.0	GR_A	93	0.91

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
						GS_A	145	0.00
	4,302,305	T	G		70.0	GR_A	70	1.00
						GS_A	149	0.00
	4,302,311	TGGGGAT	TGGGAT		106.0	GR_A	106	1.00
						GS_A	143	0.00
10	636,470	ATGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTTA	ATGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTTA	del	1,449.0	GR_A	160	1.00
						GS_A	160	0.05
	936,779	CGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GC	CGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGC	ins	510.0	GR_A	146	1.00
						GS_A	146	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	953,866	TTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA		586.0	GR_C	146	1.00
						GS_C	146	0.00
	1,085,346	AGGGGGGGGGGTTGCT	AGGGGGGGGGGGTTGCT		1,744.0	GR_C	160	0.88
						GS_C	160	0.00
	2,376,963	TACCACCACCACCACCACCACACCACCACCACCACCACCACCACCACCACCAG	TACCACCACCACCACCACCACACCACCACCACCACCACCACCACCACCACCACCAG		171.0	GR_A	138	0.91
						GS_A	160	0.00
	4,120,666	GCACACACACACACACACACG	GCACACACACACACACACACACG		2,024.0	GR_A	145	0.97
						GS_A	145	0.00
	4,444,550	CGA	CGATGTTAACGAAGTGA		2,085.0	GR_A	138	0.86
						GS_A	138	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	4,908,592	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTTTGT GTGTT	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTTT GTGTGTT		283.0	GR_B	160	1.00
						GS_B	160	0.00
	5,098,883	TGGGGGGGGG GCAGGGCCA	TGGGGGGGGG GGCAGGGCCA		184.0	GR_A	142	0.91
						GS_A	142	0.00
	5,165,203	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTTG	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTTG		342.0	GR_A	145	1.00
						GS_A	145	0.00
	5,704,537	TGGCGGCGGC GGCGGCGGCG GCGGCGGCGG CGGCGGCGGC GGCGGCGGCG GCGGCGGCCC GGGCGCT	TGGCGGCGGC GGCGGCGGCG GCGGCGGCGG CGGCGGCGGC GGCGGCGGCG GCGGCCCGGG CGCT	del	1,001.0	GR_B	150	1.00
						GS_B	150	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	5,738,864	GGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTTTA	GGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTTTA	ins	908.0	GR_B	160	1.00
						GS_B	160	0.00
	5,823,824	CGTGTGTGTGTGTGTGTGTGTA	CGTGTGTGTGTGTGTGTA	del	524.0	GR_A	148	1.00
						GS_A	148	0.00
	6,084,995	AACACACACACACACACACACACACACACACACG	AACACACACACACACACACACACACACACACG	ins	762.0	GR_C	160	0.94
						GS_C	160	0.00
	6,576,861	GGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC	GGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC		310.0	GR_A	143	0.94
						GS_A	143	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	6,618,841	TACACACACACACACACACAT	TACACACACACACACAT	del	1,054.0	GR_A	160	1.00
						GS_A	160	0.00
	6,719,353	GTTTAG	GTTTTAG	ins	350.0	GR_B	142	1.00
						GS_B	125	0.27
11	75,792	AGTGTGTGTGTGTGTGTGTGTGA	AGTGTGTGTGTGTGTGTGA	del	1,313.0	GR_A	152	1.00
						GS_A	152	0.02
	3,485,551	CCACACACACACACACACACAAG	CCACACACACACACACAAG		1,174.0	GR_A	142	0.98
						GS_A	142	0.00
12	942,978	CCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCC	CCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGCCTCCCTCGC	ins	2,132.0	GR_B	160	0.88
						GS_B	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	2,719,003	A	G	snp	3,349.0	GR_A	160	1.00
						GS_A	160	0.00
	9,028,933	C	T		4,732.0	GR_A	149	1.00
						GS_A	149	0.00
13	2,485,352	AACACACACACACACACACT	AACACACACACACACACACT	ins	2,759.0	GR_A	160	1.00
						GS_A	160	0.00
	2,897,208	AGCCGCCGCCGCCA	AGCCGCCGCCGCCA		2,993.0	GR_A	160	1.00
						GS_A	160	0.00
	3,432,303	TGGCGGCGGCGGCGGCGGCT	TGGCGGCGGCGGCGGCGGCT		2,402.0	GR_A	148	1.00
						GS_A	148	0.00
14	115,361	CCTGCTGCTGCTGCTGCTGCTGCTGCTGC	CCTGCTGCTGCTGCTGCTGCTGCTGCTGC		2,223.0	GR_A	157	0.95
						GS_A	157	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	132,963	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTTGTC	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTTGTC		249.0	GR_A	160	0.86
						GS_A	160	0.00
	148,097	C	T	snp	3,678.0	GR_A	160	0.99
						GS_A	160	0.00
	379,694	TCACACACACACACACACACACACAAACACACACG	TCACACACACACACACACACACACACAAACACACACG	ins	211.0	GR_A	160	0.92
						GS_A	160	0.00
	425,998	AGAGGAGGAGGAGGAGGAGGAGGAGGT	AGAGGAGGAGGAGGAGGAGGAGGAGGAGGT		438.0	GR_A	160	1.00
						GS_A	160	0.00
	3,450,053				940.0	GR_A	160	0.95

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		TACACACACACACACACACACACACACACACAAG	TACACACACACACACACACACACACACACAAG			GS_A	160	0.00
15	555,528	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTTA	del	267.0	GR_A	140	1.00
						GS_A	140	0.00
	600,561	TACACACACACACACACACACACACACG	TACACACACACACACACACACACACG		440.0	GR_A	160	0.90
						GS_A	160	0.00
	634,759	AACACACACACACACACACACACACACACACACA	AACACACACACACACACACACACACACACACACA	ins	213.0	GR_C	142	1.00
						GS_C	142	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		CACACACACACACACACACG	CACACACACACG					
	1,730,069	CGTGTGTGTGTGTGTGTA	CGTGTGTGTGTGTGTGTA		266.0	GR_B	160	1.00
						GS_B	160	0.00
16	311	C	A	snp	552.0	GR_C	140	0.96
						GS_C	140	0.11
	207,851	GCACACACACACACACACACACACACACACAT	GCACACACACACACACACACACACACACAT		883.0	GR_B	136	1.00
						GS_B	136	0.00
	2,784,507	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	ins	278.0	GR_A	160	0.89
						GS_A	160	0.00
	3,700,607				2,095.0	GR_A	146	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		TGGCGCGGGGCGCGGGCGCGGGGA	TGGCGCGGGGCGCGGGCGCGGGCGCGGGGA			GS_A	146	0.00
	3,814,125	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	del	365.0	GR_B	150	1.00
						GS_B	150	0.00
17	280,634	AGGGGGGGGGCGCT	AGGGGGGGGGGCGCT	ins	4,014.0	GR_A	160	1.00
						GS_A	160	0.00
	999,908	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGC		640.0	GR_C	160	0.93
						GS_C	160	0.00
	1,358,051				1,463.0	GR_A	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
		ATGTGTGTGTGTGTGTGTGTGTGC	ATGTGTGTGTGTGTGTGTGTGTGC			GS_A	160	0.00
	3,833,821	GCACACACACACACACACAAACACACACG	GCACACACACACACACACAAACACACACG		1,597.0	GR_B	160	1.00
						GS_B	160	0.00
	4,756,982	CCACACACACACACACACAT	CCACACACACACACACACAT		354.0	GR_A	160	1.00
						GS_A	160	0.00
	4,812,973	TCACACACACACACACACACACACAT	TCACACACACACACACACACACACAT		273.0	GR_A	160	1.00
						GS_A	160	0.00
	4,925,673	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT		3,328.0	GR_A	160	0.96
						GS_A	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Sample	GQ	AAF
	5,995,763	GTTTCTTTTTCTG	GTTTTCTTTATTTTCTG	complex	701.0	GR_A	145	1.00
						GS_A	145	0.00
	5,995,777	CA	TG		852.0	GR_A	145	1.00
						GS_A	145	0.00
18	1,289,528	CTATATATATATAAG	CTATATATATAAG	del	2,213.0	GR_A	152	1.00
						GS_A	152	0.00
	1,403,655	TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCACC	TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCACC		912.0	GR_A	140	1.00
						GS_A	140	0.00

Appendix 4: Annotation of Variants susceptible to support glyphosate resistance from the SR experiment (Chapter3)

SnpEff (v4.3+T.galaxy2) Annotation of variants susceptible to support glyphosate resistance that are reported in Appendix 3

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
05	3,685,946	Cre05.g233702-CHR_END	3	intergenic region	MODIFIER
01	19,761	Cre01.g000017	2	upstream gene variant	MODIFIER
		Cre01.g000033	2	upstream gene variant	MODIFIER
		CHR_START-Cre01.g000017	2	intergenic region	MODIFIER
	1,555,693	Cre01.g008450	2	upstream gene variant	MODIFIER
		Cre01.g008500	2	intron variant	MODIFIER
	2,390,845	Cre01.g013750	2	upstream gene variant	MODIFIER
		Cre01.g013769	2	upstream gene variant	MODIFIER
		Cre01.g013801	2	upstream gene variant	MODIFIER
		Cre01.g013700	2	downstream gene variant	MODIFIER
		Cre01.g013700	2	downstream gene variant	MODIFIER
		Cre01.g013800	2	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	3,980,468	Cre01.g026350	2	frameshift variant	HIGH
		Cre01.g026400	2	upstream gene variant	MODIFIER
		Cre01.g026300	2	downstream gene variant	MODIFIER
	4,097,932	Cre01.g027450	2	3 prime UTR variant	MODIFIER
		Cre01.g027500	2	upstream gene variant	MODIFIER
		Cre01.g027400	2	downstream gene variant	MODIFIER
		Cre01.g027550	2	downstream gene variant	MODIFIER
		Cre01.g027400	2	downstream gene variant	MODIFIER
	4,740,525	Cre01.g033250	2	disruptive inframe insertion	MODERATE
		Cre01.g033200	2	upstream gene variant	MODIFIER
		Cre01.g033300	2	upstream gene variant	MODIFIER
		Cre01.g033300	2	upstream gene variant	MODIFIER
	5,393,698	Cre01.g038400	2	missense variant	MODERATE
		Cre01.g038450	2	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
02	6,475,904	Cre01.g046237	2	upstream gene variant	MODIFIER
		Cre01.g046237	2	upstream gene variant	MODIFIER
		Cre01.g046324	2	intron variant	MODIFIER
	7,148,825	Cre01.g051625	2	downstream gene variant	MODIFIER
		Cre01.g051625-Cre01.g051700	2	intergenic region	MODIFIER
	7,304,996	Cre01.g051750	2	3 prime UTR variant	MODIFIER
		Cre01.g051750	2	downstream gene variant	MODIFIER
		Cre01.g051800	2	downstream gene variant	MODIFIER
	420,383	Cre02.g076000	2	upstream gene variant	MODIFIER
		Cre02.g076100	2	downstream gene variant	MODIFIER
		Cre02.g076050	2	intron variant	MODIFIER
	1,019,958	Cre02.g080300	2	missense variant	MODERATE
		Cre02.g080350	2	upstream gene variant	MODIFIER
		Cre02.g080250	2	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	3,116,800	Cre02.g096650	2	downstream gene variant	MODIFIER
		Cre02.g096700	2	intron variant	MODIFIER
	3,567,631	Cre02.g099950	2	3 prime UTR variant	MODIFIER
		Cre02.g099900	2	upstream gene variant	MODIFIER
		Cre02.g100000	2	downstream gene variant	MODIFIER
	5,354,377	Cre02.g114000	2	upstream gene variant	MODIFIER
		Cre02.g114050	2	upstream gene variant	MODIFIER
		Cre02.g114001	2	intron variant	MODIFIER
	5,938,736	Cre02.g118700	2	intron variant	MODIFIER
	6,585,231	Cre09.g387060	2	3 prime UTR variant	MODIFIER
		Cre09.g387097	2	downstream gene variant	MODIFIER
03	4,890,271	Cre03.g179350	2	5 prime UTR variant	MODIFIER
		Cre03.g179300	2	upstream gene variant	MODIFIER
		Cre03.g179450	2	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre03.g179300	2	upstream gene variant	MODIFIER
		Cre03.g179400	2	downstream gene variant	MODIFIER
	4,939,424	Cre03.g179880	2	conservative inframe insertion	MODERATE
		Cre03.g179901	2	upstream gene variant	MODIFIER
		Cre03.g179860	2	downstream gene variant	MODIFIER
04	32	CHR_START-Cre04.g213761	2	intergenic region	MODIFIER
	82,722	Cre04.g214657	2	conservative inframe insertion	MODERATE
		Cre04.g214769	2	downstream gene variant	MODIFIER
	2,493,939	Cre04.g222450	2	downstream gene variant	MODIFIER
		Cre04.g222500	2	downstream gene variant	MODIFIER
		Cre04.g222450-Cre04.g222500	2	intergenic region	MODIFIER
	3,658,937	Cre04.g229350	2	downstream gene variant	MODIFIER
		Cre04.g229398	2	downstream gene variant	MODIFIER
		Cre04.g229374	2	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
05	3,942,575	Cre04.g232502	2	downstream gene variant	MODIFIER
		Cre04.g232402	2	intron variant	MODIFIER
	635,242	Cre05.g245150	2	frameshift variant	HIGH
		Cre05.g245100	2	upstream gene variant	MODIFIER
06	384,677	Cre06.g251450	2	upstream gene variant	MODIFIER
		Cre06.g251500	2	downstream gene variant	MODIFIER
		Cre06.g251550	2	intron variant	MODIFIER
	1,669,850	Cre06.g261050	2	downstream gene variant	MODIFIER
		Cre06.g261026	2	intron variant	MODIFIER
	2,083,864	Cre06.g264850	2	intron variant	MODIFIER
	3,385,106	Cre06.g277400	2	upstream gene variant	MODIFIER
		Cre06.g277450	2	intron variant	MODIFIER
	6,463,824	Cre06.g293400	2	upstream gene variant	MODIFIER
		Cre06.g293450	2	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre06.g293450	2	intron variant	MODIFIER
	7,000,602	Cre06.g297082	2	conservative inframe insertion	MODERATE
		Cre06.g297082	2	upstream gene variant	MODIFIER
	8,266,021	Cre06.g306950	2	downstream gene variant	MODIFIER
		Cre06.g306900	2	intron variant	MODIFIER
07	914,701	Cre07.g319000	2	downstream gene variant	MODIFIER
		Cre07.g318950	2	intron variant	MODIFIER
	1,010,355	Cre07.g319600	2	conservative inframe insertion	MODERATE
		Cre07.g319650	2	downstream gene variant	MODIFIER
	1,781,173	Cre07.g325727	2	conservative inframe insertion	MODERATE
		Cre07.g325728	2	upstream gene variant	MODIFIER
	3,860,633	Cre07.g339104	2	intron variant	MODIFIER
		Cre07.g339104	2	intron variant	MODIFIER
	4,678,848	Cre07.g345250	2	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre07.g345350	2	upstream gene variant	MODIFIER
		Cre07.g345300	2	intron variant	MODIFIER
08	340,984	Cre08.g358575	2	missense variant	MODERATE
	415,616	Cre08.g358650	2	upstream gene variant	MODIFIER
		Cre08.g358616	2	downstream gene variant	MODIFIER
		Cre08.g358600	2	intron variant	MODIFIER
	3,215,241	Cre08.g376740	2	upstream gene variant	MODIFIER
		Cre08.g376740	2	upstream gene variant	MODIFIER
		Cre08.g376720	2	intron variant	MODIFIER
		Cre08.g376720	2	intron variant	MODIFIER
	3,918,420	Cre08.g382050	2	upstream gene variant	MODIFIER
		Cre08.g381950	2	downstream gene variant	MODIFIER
		Cre08.g381983	2	intron variant	MODIFIER
		Cre08.g381983	2	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	4,439,267	Cre08.g385350	2	upstream gene variant	MODIFIER
		Cre08.g385400	2	upstream gene variant	MODIFIER
		Cre08.g385400	2	upstream gene variant	MODIFIER
		Cre08.g385500	2	downstream gene variant	MODIFIER
		Cre08.g385500	2	downstream gene variant	MODIFIER
		Cre08.g385450	2	intron variant	MODIFIER
09	810,582	Cre09.g403800	2	missense variant	MODERATE
	1,439,622	Cre09.g399908	2	downstream gene variant	MODIFIER
		Cre09.g399907	2	intron variant	MODIFIER
	2,467,744	Cre09.g393150	2	intron variant	MODIFIER
	5,638,715	Cre09.g410100	2	downstream gene variant	MODIFIER
		Cre09.g410200	2	downstream gene variant	MODIFIER
		Cre09.g410150	2	intron variant	MODIFIER
	6,380,673	Cre09.g414800	2	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre09.g414900	2	downstream gene variant	MODIFIER
		Cre09.g414800	2	downstream gene variant	MODIFIER
		Cre09.g414850	2	intron variant	MODIFIER
		Cre09.g414850	2	intron variant	MODIFIER
10	900,763	Cre10.g423750	2	downstream gene variant	MODIFIER
		Cre10.g423800	2	intron variant	MODIFIER
	2,116,960	Cre10.g433200	2	downstream gene variant	MODIFIER
		Cre10.g433150	2	intron variant	MODIFIER
	4,436,009	Cre10.g450700	2	upstream gene variant	MODIFIER
		Cre10.g450626	2	downstream gene variant	MODIFIER
		Cre10.g450650	2	intron variant	MODIFIER
	4,528,327	Cre10.g451400	2	intron variant	MODIFIER
	6,215,827	Cre10.g463400	2	upstream gene variant	MODIFIER
		Cre10.g463450	2	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre10.g463500	2	intron variant	MODIFIER
	6,518,000	Cre10.g465700	2	upstream gene variant	MODIFIER
		Cre10.g465650	2	intron variant	MODIFIER
	6,719,219	Cre10.g467200-CHR_END	2	intergenic region	MODIFIER
11	1,039,723	Cre11.g467669	2	intron variant	MODIFIER
	3,189,268	Cre11.g475650	2	disruptive inframe insertion	MODERATE
		Cre11.g475600	2	downstream gene variant	MODIFIER
		Cre11.g475626	2	downstream gene variant	MODIFIER
		Cre11.g475700	2	downstream gene variant	MODIFIER
12	1,296,987	Cre12.g489300	2	intron variant	MODIFIER
	1,535,558	Cre12.g487450	2	upstream gene variant	MODIFIER
		Cre12.g487402	2	upstream gene variant	MODIFIER
		Cre12.g487400	2	downstream gene variant	MODIFIER
		Cre12.g487350	2	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre12.g487400-Cre12.g487402	2	intergenic region	MODIFIER
	7,880,011	Cre12.g553350	2	upstream gene variant	MODIFIER
		Cre12.g553250	2	upstream gene variant	MODIFIER
		Cre16.g654100	2	downstream gene variant	MODIFIER
		Cre12.g553302	2	downstream gene variant	MODIFIER
		Cre16.g654100	2	downstream gene variant	MODIFIER
		Cre12.g553300	2	intron variant	MODIFIER
13	1,266,357	Cre13.g570951	2	upstream gene variant	MODIFIER
		Cre13.g570851	2	downstream gene variant	MODIFIER
		Cre13.g570900	2	intron variant	MODIFIER
	3,411,519	Cre13.g586950	2	upstream gene variant	MODIFIER
		Cre13.g587050	2	downstream gene variant	MODIFIER
		Cre13.g587000	2	intron variant	MODIFIER
	4,149,461	Cre13.g591350	2	3 prime UTR variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre13.g591300	2	upstream gene variant	MODIFIER
		Cre13.g591400	2	upstream gene variant	MODIFIER
	4,798,082	Cre13.g604850	2	downstream gene variant	MODIFIER
		Cre13.g604950	2	downstream gene variant	MODIFIER
		Cre13.g604950	2	downstream gene variant	MODIFIER
		Cre13.g604950	2	downstream gene variant	MODIFIER
		Cre13.g604905	2	intron variant	MODIFIER
14	422,685	Cre14.g610700	2	downstream gene variant	MODIFIER
		Cre14.g610663	2	intron variant	MODIFIER
	1,532,630	Cre14.g618400	2	upstream gene variant	MODIFIER
		Cre14.g618450	2	upstream gene variant	MODIFIER
		Cre14.g618400-Cre14.g618450	2	intergenic region	MODIFIER
	2,276,825	Cre14.g623050	2	upstream gene variant	MODIFIER
		Cre14.g623050	2	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre14.g622951	2	downstream gene variant	MODIFIER
		Cre14.g623000	2	intron variant	MODIFIER
	3,850,348	Cre14.g632759	2	upstream gene variant	MODIFIER
		Cre14.g632775	2	downstream gene variant	MODIFIER
		Cre14.g632767	2	intron variant	MODIFIER
	4,186,933	Cre14.g634322-CHR_END	2	intergenic region	MODIFIER
	4,188,448	Cre14.g634322-CHR_END	2	intergenic region	MODIFIER
	4,188,777	Cre14.g634322-CHR_END	2	intergenic region	MODIFIER
15	3,555,493	Cre23.g754897	2	intron variant	MODIFIER
16	512,172	Cre16.g692585	2	downstream gene variant	MODIFIER
		Cre16.g692600	2	intron variant	MODIFIER
		Cre16.g692600	2	intron variant	MODIFIER
	2,260,411	Cre16.g658950	2	conservative inframe insertion	MODERATE
		Cre16.g659000	2	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre16.g658926	2	downstream gene variant	MODIFIER
	2,815,045	Cre16.g663200	2	downstream gene variant	MODIFIER
		Cre16.g663280	2	downstream gene variant	MODIFIER
		Cre16.g663250	2	intron variant	MODIFIER
		Cre16.g663250	2	intron variant	MODIFIER
	4,948,956	Cre16.g685800	2	downstream gene variant	MODIFIER
		Cre16.g685901	2	intron variant	MODIFIER
	5,230,480	Cre16.g683300	2	3 prime UTR variant	MODIFIER
		Cre16.g683250	2	upstream gene variant	MODIFIER
		Cre16.g683350	2	downstream gene variant	MODIFIER
		Cre16.g683200	2	downstream gene variant	MODIFIER
	7,716,527	Cre16.g689759	2	downstream gene variant	MODIFIER
		Cre16.g689647	2	intron variant	MODIFIER
17	2,372,299	Cre17.g715176	2	3 prime UTR variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre17.g715200	2	upstream gene variant	MODIFIER
		Cre17.g715100	2	downstream gene variant	MODIFIER
		Cre17.g715150	2	downstream gene variant	MODIFIER
	4,182,600	Cre17.g729950	2	upstream gene variant	MODIFIER
		Cre17.g730050	2	upstream gene variant	MODIFIER
		Cre17.g730000	2	intron variant	MODIFIER
	5,020,066	Cre17.g736350	2	conservative inframe deletion	MODERATE
		Cre17.g736329	2	downstream gene variant	MODIFIER
		Cre17.g736400	2	downstream gene variant	MODIFIER
	6,471,328	Cre17.g744797	2	missense variant&conservative inframe insertion	MODERATE
		Cre17.g744747	2	upstream gene variant	MODIFIER
		Cre17.g744747	2	upstream gene variant	MODIFIER
01	564,397	Cre01.g003475	1	splice region variant&intron variant	LOW

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre01.g003463	1	downstream gene variant	MODIFIER
		Cre01.g003487	1	downstream gene variant	MODIFIER
	2,397,527	Cre01.g013800	1	downstream gene variant	MODIFIER
		Cre01.g013850	1	downstream gene variant	MODIFIER
		Cre01.g013900	1	downstream gene variant	MODIFIER
		Cre01.g013900	1	downstream gene variant	MODIFIER
		Cre01.g013801	1	intron variant	MODIFIER
	3,152,613	Cre01.g020182	1	upstream gene variant	MODIFIER
		Cre01.g020223	1	downstream gene variant	MODIFIER
		Cre01.g020182-Cre01.g020223	1	intergenic region	MODIFIER
	3,252,015	Cre01.g020950	1	upstream gene variant	MODIFIER
		Cre01.g020918	1	intron variant	MODIFIER
	4,839,802	Cre01.g034000	1	upstream gene variant	MODIFIER
		Cre01.g034100	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre01.g034100	1	upstream gene variant	MODIFIER
		Cre01.g034050	1	intron variant	MODIFIER
	5,621,006	Cre01.g040517	1	upstream gene variant	MODIFIER
		Cre01.g040533	1	upstream gene variant	MODIFIER
		Cre01.g040500	1	downstream gene variant	MODIFIER
		Cre01.g040550	1	downstream gene variant	MODIFIER
		Cre01.g040517-Cre01.g040533	1	intergenic region	MODIFIER
02	423,838	Cre02.g076000	1	upstream gene variant	MODIFIER
		Cre02.g076100	1	downstream gene variant	MODIFIER
		Cre02.g076050	1	intron variant	MODIFIER
	994,394	Cre02.g080050	1	upstream gene variant	MODIFIER
		Cre02.g080100	1	upstream gene variant	MODIFIER
		Cre02.g080050-Cre02.g080100	1	intergenic region	MODIFIER
	1,053,941	Cre02.g080500	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre02.g080600	1	downstream gene variant	MODIFIER
		Cre02.g080550	1	intron variant	MODIFIER
	1,835,505	Cre02.g087600	1	downstream gene variant	MODIFIER
		Cre02.g087551	1	intron variant	MODIFIER
	5,399,695	Cre02.g114250	1	upstream gene variant	MODIFIER
		Cre02.g114350	1	downstream gene variant	MODIFIER
		Cre02.g114300	1	intron variant	MODIFIER
	6,249,627	Cre09.g387650	1	downstream gene variant	MODIFIER
		Cre09.g387750	1	downstream gene variant	MODIFIER
		Cre09.g387700	1	intron variant	MODIFIER
	6,857,626	Cre09.g389060	1	upstream gene variant	MODIFIER
		Cre09.g388986	1	downstream gene variant	MODIFIER
		Cre09.g389023	1	intron variant	MODIFIER
	8,514,074	Cre02.g141050	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre02.g141000	1	intron variant	MODIFIER
03	1,385,028	Cre03.g150900	1	3 prime UTR variant	MODIFIER
		Cre03.g151000	1	upstream gene variant	MODIFIER
		Cre03.g150950	1	downstream gene variant	MODIFIER
	6,245,369	Cre03.g192501	1	conservative inframe insertion	MODERATE
		Cre03.g192450	1	downstream gene variant	MODIFIER
		Cre03.g192550	1	downstream gene variant	MODIFIER
	8,524,378	Cre03.g208833	1	upstream gene variant	MODIFIER
		Cre03.g208721	1	intron variant	MODIFIER
04	189,699	Cre04.g217000	1	conservative inframe insertion	MODERATE
	286,847	Cre04.g217750	1	intron variant	MODIFIER
	286,887	Cre04.g217750	1	intron variant	MODIFIER
	839,723	Cre04.g217976	1	upstream gene variant	MODIFIER
		Cre04.g217975-Cre04.g217976	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	1,267,796	Cre04.g214800	1	upstream gene variant	MODIFIER
		Cre04.g214801	1	downstream gene variant	MODIFIER
		Cre04.g214700	1	downstream gene variant	MODIFIER
		Cre04.g214801	1	downstream gene variant	MODIFIER
		Cre04.g214801	1	downstream gene variant	MODIFIER
		Cre04.g214750	1	intron variant	MODIFIER
	1,397,506	Cre04.g214351	1	downstream gene variant	MODIFIER
		Cre04.g214250	1	intron variant	MODIFIER
		Cre04.g214250	1	intron variant	MODIFIER
	1,863,043	Cre04.g218350	1	upstream gene variant	MODIFIER
		Cre04.g218350	1	upstream gene variant	MODIFIER
		Cre04.g218300	1	intron variant	MODIFIER
	2,087,427	Cre04.g219750	1	intron variant	MODIFIER
	2,172,588	Cre04.g220100	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre04.g220200	1	downstream gene variant	MODIFIER
		Cre04.g220200	1	downstream gene variant	MODIFIER
		Cre04.g220200	1	downstream gene variant	MODIFIER
		Cre04.g220150	1	intron variant	MODIFIER
	3,428,879	Cre04.g228100	1	upstream gene variant	MODIFIER
		Cre04.g228050	1	downstream gene variant	MODIFIER
		Cre04.g228000	1	intron variant	MODIFIER
05	48,804	Cre05.g241751	1	downstream gene variant	MODIFIER
		Cre05.g241750	1	intron variant	MODIFIER
	433,737	Cre05.g243801	1	intron variant	MODIFIER
		Cre05.g243801	1	intron variant	MODIFIER
		Cre05.g243801	1	intron variant	MODIFIER
	685,392	Cre05.g245351	1	upstream gene variant	MODIFIER
		Cre05.g245352	1	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre05.g245352	1	intron variant	MODIFIER
	922,926	Cre05.g246753	1	conservative inframe insertion	MODERATE
		Cre05.g246650	1	downstream gene variant	MODIFIER
		Cre05.g246752	1	downstream gene variant	MODIFIER
	1,048,297	Cre05.g247650	1	intron variant	MODIFIER
	1,129,258	Cre05.g248250	1	upstream gene variant	MODIFIER
		Cre05.g248300	1	intron variant	MODIFIER
	1,654,763	Cre05.g235186-Cre10.g445299	1	intergenic region	MODIFIER
	2,880,401	Cre24.g755297	1	upstream gene variant	MODIFIER
		Cre24.g755347	1	intron variant	MODIFIER
06	1,413,415	Cre06.g258950	1	disruptive inframe insertion	MODERATE
		Cre06.g258900	1	downstream gene variant	MODIFIER
		Cre06.g259000	1	downstream gene variant	MODIFIER
	1,471,155	Cre06.g259401	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre06.g259476	1	downstream gene variant	MODIFIER
		Cre06.g259500	1	downstream gene variant	MODIFIER
		Cre06.g259450	1	intron variant	MODIFIER
	2,084,732	Cre06.g264850	1	intron variant	MODIFIER
	3,699,555	Cre06.g278135	1	upstream gene variant	MODIFIER
		Cre06.g278137	1	upstream gene variant	MODIFIER
		Cre06.g278138	1	downstream gene variant	MODIFIER
		Cre06.g278136	1	intron variant	MODIFIER
	4,152,923	Cre06.g278225	1	upstream gene variant	MODIFIER
		Cre06.g278221	1	downstream gene variant	MODIFIER
		Cre06.g278222	1	downstream gene variant	MODIFIER
		Cre06.g278224	1	downstream gene variant	MODIFIER
		Cre06.g278223	1	intron variant	MODIFIER
	4,726,506	Cre06.g278650	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre06.g278750	1	downstream gene variant	MODIFIER
		Cre06.g278700	1	intron variant	MODIFIER
	5,210,759	Cre06.g283450	1	missense variant	MODERATE
		Cre06.g283400	1	upstream gene variant	MODIFIER
		Cre06.g283500	1	downstream gene variant	MODIFIER
	6,441,214	Cre06.g293300	1	disruptive inframe insertion	MODERATE
		Cre06.g293350	1	downstream gene variant	MODIFIER
	6,704,670	Cre06.g295050	1	upstream gene variant	MODIFIER
		Cre06.g295001	1	intron variant	MODIFIER
	7,342,788	Cre06.g299800	1	upstream gene variant	MODIFIER
		Cre06.g299900	1	upstream gene variant	MODIFIER
		Cre06.g299850	1	downstream gene variant	MODIFIER
		Cre06.g299850-Cre06.g299900	1	intergenic region	MODIFIER
	8,621,457	Cre06.g309900	1	5 prime UTR premature start codon gain variant	LOW

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre06.g309900	1	5 prime UTR variant	MODIFIER
		Cre06.g309951	1	downstream gene variant	MODIFIER
07	431,386	Cre07.g315200	1	missense variant	MODERATE
		Cre07.g315100	1	upstream gene variant	MODIFIER
		Cre07.g315150	1	upstream gene variant	MODIFIER
	3,582,932	Cre07.g337000	1	upstream gene variant	MODIFIER
		Cre07.g337100	1	downstream gene variant	MODIFIER
		Cre07.g337050	1	intron variant	MODIFIER
	3,618,668	Cre07.g337400	1	upstream gene variant	MODIFIER
		Cre07.g337350	1	downstream gene variant	MODIFIER
		Cre07.g337516	1	downstream gene variant	MODIFIER
		Cre07.g337450	1	intron variant	MODIFIER
	4,133,815	Cre07.g342350	1	synonymous variant	LOW
		Cre07.g342250	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre07.g342402	1	downstream gene variant	MODIFIER
		Cre07.g342352	1	downstream gene variant	MODIFIER
	4,166,695	Cre07.g342052	1	upstream gene variant	MODIFIER
		Cre07.g341950	1	upstream gene variant	MODIFIER
		Cre07.g342000	1	intron variant	MODIFIER
	4,729,476	Cre07.g345700	1	upstream gene variant	MODIFIER
		Cre07.g345800	1	downstream gene variant	MODIFIER
		Cre07.g345850	1	downstream gene variant	MODIFIER
		Cre07.g345750	1	intron variant	MODIFIER
	4,769,875	Cre07.g346000	1	missense variant	MODERATE
	5,053,746	Cre07.g348010	1	upstream gene variant	MODIFIER
		Cre07.g347980	1	intron variant	MODIFIER
		Cre07.g347980	1	intron variant	MODIFIER
	6,126,432	Cre07.g356200	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre07.g356250	1	intron variant	MODIFIER
08	286,819	Cre08.g358570	1	upstream gene variant	MODIFIER
		Cre08.g358569	1	intron variant	MODIFIER
	767,586	Cre08.g360650	1	downstream gene variant	MODIFIER
		Cre08.g360600	1	intron variant	MODIFIER
	1,305,433	Cre08.g363800	1	upstream gene variant	MODIFIER
		Cre08.g363837	1	intron variant	MODIFIER
	1,516,859	Cre08.g364800	1	downstream gene variant	MODIFIER
		Cre08.g364850	1	intron variant	MODIFIER
	1,621,824	Cre08.g365664	1	upstream gene variant	MODIFIER
		Cre08.g365720	1	downstream gene variant	MODIFIER
		Cre08.g365692	1	intron variant	MODIFIER
	1,774,937	Cre08.g366579	1	conservative inframe insertion	MODERATE
		Cre08.g366550	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre08.g366600	1	upstream gene variant	MODIFIER
	3,012,012	Cre08.g374950	1	splice acceptor variant&intron variant	HIGH
		Cre08.g375000	1	upstream gene variant	MODIFIER
	1,183,240	Cre09.g401050	1	intron variant	MODIFIER
	3,186,690	Cre09.g394621	1	intron variant	MODIFIER
		Cre09.g394621	1	intron variant	MODIFIER
	3,621,877	Cre09.g396994	1	downstream gene variant	MODIFIER
		Cre09.g397068	1	downstream gene variant	MODIFIER
09		Cre09.g397031	1	intron variant	MODIFIER
	3,801,124	Cre09.g397771	1	disruptive inframe insertion	MODERATE
	4,096,128	Cre09.g399289	1	3 prime UTR variant	MODIFIER
		Cre09.g399363	1	upstream gene variant	MODIFIER
		Cre09.g399326	1	downstream gene variant	MODIFIER
	4,302,169	Cre09.g400034	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre09.g400034-Cre09.g400071	1	intergenic region	MODIFIER
	4,302,305	Cre09.g400034	1	upstream gene variant	MODIFIER
		Cre09.g400034-Cre09.g400071	1	intergenic region	MODIFIER
	4,302,311	Cre09.g400034	1	upstream gene variant	MODIFIER
		Cre09.g400034-Cre09.g400071	1	intergenic region	MODIFIER
10	636,470	Cre10.g421850	1	3 prime UTR variant	MODIFIER
		Cre10.g421800	1	downstream gene variant	MODIFIER
		Cre10.g421900	1	downstream gene variant	MODIFIER
		Cre10.g421800	1	downstream gene variant	MODIFIER
	936,779	Cre10.g424100	1	upstream gene variant	MODIFIER
		Cre10.g424150	1	downstream gene variant	MODIFIER
		Cre10.g424250	1	downstream gene variant	MODIFIER
		Cre10.g424200	1	intron variant	MODIFIER
	953,866	Cre10.g424400	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre10.g424300	1	downstream gene variant	MODIFIER
		Cre10.g424450	1	downstream gene variant	MODIFIER
		Cre10.g424350	1	intron variant	MODIFIER
	1,085,346	Cre10.g425200	1	3 prime UTR variant	MODIFIER
		Cre10.g425251	1	downstream gene variant	MODIFIER
	2,376,963	Cre10.g434850	1	upstream gene variant	MODIFIER
		Cre10.g434950	1	upstream gene variant	MODIFIER
		Cre10.g434900	1	intron variant	MODIFIER
	4,120,666	Cre10.g448600	1	downstream gene variant	MODIFIER
		Cre10.g448700	1	downstream gene variant	MODIFIER
		Cre10.g448650	1	intron variant	MODIFIER
	4,444,550	Cre10.g450800	1	upstream gene variant	MODIFIER
		Cre10.g450650	1	downstream gene variant	MODIFIER
		Cre10.g450700	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre10.g450750	1	downstream gene variant	MODIFIER
		Cre10.g450850	1	downstream gene variant	MODIFIER
		Cre10.g450750-Cre10.g450800	1	intergenic region	MODIFIER
	4,908,592	Cre10.g453782	1	downstream gene variant	MODIFIER
		Cre10.g453800	1	downstream gene variant	MODIFIER
		Cre10.g453807	1	intron variant	MODIFIER
		Cre10.g453807	1	intron variant	MODIFIER
	5,098,883	Cre10.g455190	1	3 prime UTR variant	MODIFIER
		Cre10.g455231	1	upstream gene variant	MODIFIER
	5,165,203	Cre10.g455700	1	upstream gene variant	MODIFIER
		Cre10.g455750	1	intron variant	MODIFIER
	5,704,537	Cre10.g459700	1	conservative inframe deletion	MODERATE
		Cre10.g459750	1	upstream gene variant	MODIFIER
		Cre10.g459800	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	5,738,864	Cre10.g460050	1	downstream gene variant	MODIFIER
		Cre10.g460050	1	downstream gene variant	MODIFIER
		Cre10.g460100	1	intron variant	MODIFIER
	5,823,824	Cre10.g460600	1	upstream gene variant	MODIFIER
		Cre10.g460700	1	upstream gene variant	MODIFIER
		Cre10.g460650	1	intron variant	MODIFIER
	6,084,995	Cre10.g462350	1	upstream gene variant	MODIFIER
		Cre10.g462300	1	intron variant	MODIFIER
	6,576,861	Cre10.g466150	1	disruptive inframe insertion	MODERATE
		Cre10.g466100	1	upstream gene variant	MODIFIER
		Cre10.g466175	1	upstream gene variant	MODIFIER
	6,618,841	Cre10.g466350	1	upstream gene variant	MODIFIER
		Cre10.g466450	1	upstream gene variant	MODIFIER
		Cre10.g466400	1	intron variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre10.g466400	1	intron variant	MODIFIER
	6,719,353	Cre10.g467200-CHR_END	1	intergenic region	MODIFIER
11	75,792	Cre11.g467529	1	intron variant	MODIFIER
	3,485,551	Cre11.g477400	1	intron variant	MODIFIER
12	942,978	Cre12.g495959-Cre12.g491950	1	intergenic region	MODIFIER
	2,719,003	Cre12.g504100	1	upstream gene variant	MODIFIER
		Cre12.g504000	1	downstream gene variant	MODIFIER
		Cre12.g504050	1	intron variant	MODIFIER
	9,028,933	Cre12.g543550	1	5 prime UTR variant	MODIFIER
		Cre12.g543500	1	upstream gene variant	MODIFIER
		Cre12.g543650	1	downstream gene variant	MODIFIER
13	2,485,352	Cre13.g579950	1	upstream gene variant	MODIFIER
		Cre13.g579901	1	intron variant	MODIFIER
	2,897,208	Cre13.g583500	1	conservative inframe insertion	MODERATE

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre13.g583450	1	downstream gene variant	MODIFIER
	3,432,303	Cre13.g587200	1	disruptive inframe insertion	MODERATE
		Cre13.g587150	1	downstream gene variant	MODIFIER
		Cre13.g587250	1	downstream gene variant	MODIFIER
14	115,361	Cre14.g608400	1	conservative inframe insertion	MODERATE
		Cre14.g608452	1	downstream gene variant	MODIFIER
	132,963	Cre14.g608500	1	upstream gene variant	MODIFIER
		Cre14.g608550	1	upstream gene variant	MODIFIER
		Cre14.g608452	1	intron variant	MODIFIER
	148,097	Cre14.g608652	1	3 prime UTR variant	MODIFIER
		Cre14.g608600	1	downstream gene variant	MODIFIER
		Cre14.g608700	1	downstream gene variant	MODIFIER
	379,694	Cre14.g610501	1	3 prime UTR variant	MODIFIER
		Cre14.g610450	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
	425,998	Cre14.g610700	1	conservative inframe insertion	MODERATE
		Cre14.g610750	1	upstream gene variant	MODIFIER
		Cre14.g610663	1	downstream gene variant	MODIFIER
	3,450,053	Cre14.g630750	1	intron variant	MODIFIER
15	555,528	Cre15.g636450	1	upstream gene variant	MODIFIER
		Cre15.g636350	1	downstream gene variant	MODIFIER
		Cre15.g636400	1	intron variant	MODIFIER
		Cre15.g636400	1	intron variant	MODIFIER
	600,561	Cre15.g636650	1	intron variant	MODIFIER
	634,759	Cre15.g636896	1	upstream gene variant	MODIFIER
		Cre15.g637000	1	upstream gene variant	MODIFIER
		Cre15.g636950	1	intron variant	MODIFIER
	1,730,069	Cre02.g141506	1	upstream gene variant	MODIFIER
		Cre19.g751347	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre19.g751347-Cre02.g141506	1	intergenic region	MODIFIER
16	311	Cre16.g653651	1	upstream gene variant	MODIFIER
		CHR_START-Cre16.g653651	1	intergenic region	MODIFIER
	207,851	Cre16.g694850	1	upstream gene variant	MODIFIER
		Cre16.g694800	1	downstream gene variant	MODIFIER
		Cre16.g694809	1	intron variant	MODIFIER
	2,784,507	Cre16.g662951	1	intron variant	MODIFIER
	3,700,607	Cre16.g686641	1	upstream gene variant	MODIFIER
		Cre16.g685929	1	downstream gene variant	MODIFIER
		Cre16.g688450	1	downstream gene variant	MODIFIER
		Cre16.g686285	1	intron variant	MODIFIER
	3,814,125	Cre16.g689201	1	downstream gene variant	MODIFIER
		Cre16.g689250	1	intron variant	MODIFIER
17	280,634	Cre17.g698100	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre17.g698150	1	upstream gene variant	MODIFIER
		Cre17.g698100-Cre17.g698150	1	intergenic region	MODIFIER
	999,908	Cre17.g703400	1	3 prime UTR variant	MODIFIER
		Cre17.g703450	1	upstream gene variant	MODIFIER
	1,358,051	Cre17.g706300	1	upstream gene variant	MODIFIER
		Cre17.g706200	1	downstream gene variant	MODIFIER
		Cre17.g706250	1	intron variant	MODIFIER
	3,833,821	Cre17.g727950	1	upstream gene variant	MODIFIER
		Cre17.g727900	1	intron variant	MODIFIER
	4,756,982	Cre17.g734821	1	upstream gene variant	MODIFIER
		Cre17.g734805	1	intron variant	MODIFIER
		Cre17.g734805	1	intron variant	MODIFIER
	4,812,973	Cre17.g735021	1	intron variant	MODIFIER
	4,925,673	Cre17.g735900	1	conservative inframe insertion	MODERATE

Chromosome	Position	Gene	Nb. replicates	Annotation	Impact score
		Cre17.g735876	1	upstream gene variant	MODIFIER
		Cre17.g735850	1	downstream gene variant	MODIFIER
	5,995,763	Cre17.g742932	1	upstream gene variant	MODIFIER
		Cre17.g742998	1	upstream gene variant	MODIFIER
		Cre17.g742866-Cre17.g742932	1	intergenic region	MODIFIER
	5,995,777	Cre17.g742932	1	upstream gene variant	MODIFIER
		Cre17.g742998	1	upstream gene variant	MODIFIER
		Cre17.g742866-Cre17.g742932	1	intergenic region	MODIFIER
18	1,289,528	Cre17.g733689	1	upstream gene variant	MODIFIER
		Cre17.g733678	1	upstream gene variant	MODIFIER
		Cre17.g733702	1	upstream gene variant	MODIFIER
		Cre17.g733650	1	intron variant	MODIFIER
	1,403,655	Cre26.g756947	1	downstream gene variant	MODIFIER
		Cre26.g756897	1	intron variant	MODIFIER

Appendix 5: List of variants susceptible to support glyphosate resistance from the VR experiment (Chapter5)

List of variants retained after filtering for high quality variants susceptible to support glyphosate resistance. For variant calling pipeline and filters details refer to Chapter 2

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
03	2,893,718	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ins	404.0	quick	GR_D	133	1.00
				ins		quick	GS_D	0	
				ins		quick	GR_E	133	1.00
				ins		quick	GS_E	0	
				ins		quick	GR_G	133	1.00
				ins		quick	GS_G	0	
				ins		interm	GR_D	133	1.00
				ins		interm	GS_D	0	
				ins		slow	GR_E	133	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				ins		slow	GS_E	0	
01	292,515	TGGGGGGGGGGTTA	TGGGGGGGGGGGTTA	ins	2,475.0	quick	GR_D	136	1.00
				ins		quick	GS_D	136	0.00
				ins		quick	GR_E	136	1.00
				ins		quick	GS_E	136	0.00
				ins		quick	GR_F	136	0.94
				ins		quick	GS_F	136	0.00
				ins		quick	GR_G	136	1.00
				ins		quick	GS_G	136	0.00
02	5,681,162	CGTGTGTGTGTGTGTGTGTGTGTGA	CGTGTGTGTGTGTGTGTGTGTGTGA	ins	9,111.0	quick	GR_D	146	1.00
				ins		quick	GS_D	146	0.00
				ins		quick	GR_E	146	1.00
				ins		quick	GS_E	146	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				ins		quick	GR_F	146	1.00
				ins		quick	GS_F	146	0.00
				ins		quick	GR_G	146	0.98
				ins		quick	GS_G	146	0.00
09	4,546,642	ACCCCCCCCCCCACACACACCTT	ACCCCCCCCCCCCACACACACCTT	ins	1,536.0	quick	GR_E	160	1.00
				ins		quick	GS_E	160	0.00
				ins		quick	GR_F	160	1.00
				ins		quick	GS_F	160	0.00
				ins		quick	GR_G	160	1.00
				ins		quick	GS_G	160	0.00
04	3,428,879	AACACACACACACACACACACACACACACACACACA	AACACACACACACACACACACACACACACACACACA	ins	384.0	interm	GR_D	137	1.00
				ins		interm	GS_D	137	0.00
				ins		slow	GR_G	137	0.91

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CACACACACA CACACACACA CACACAG	CACACACACA CACACACACA CACACACAG	ins		slow	GS_G	137	0.00
05	950,004	G	A	snp	7,496.0	interm	GR_D	160	1.00
				snp		interm	GS_D	160	0.00
				snp		interm	GR_G	160	1.00
				snp		interm	GS_G	160	0.00
07	451,071	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACCT	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACCT	ins	1,899.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
				ins		slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	668,524			del	4,179.0	interm	GR_D	145	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		AACACACACA CACACACACA CG	AACACACACA CACACACACG	del		interm	GS_D	145	0.00
				del		interm	GR_G	145	1.00
				del		interm	GS_G	145	0.01
08	581,982	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CAAT	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACAAT	ins	948.0	quick	GR_E	160	1.00
				ins		quick	GS_E	160	0.00
				ins		interm	GR_F	160	0.94
				ins		interm	GS_F	160	0.00
10	5,247,248	T	A	snp	8,247.0	interm	GR_D	152	1.00
				snp		interm	GS_D	152	0.00
				snp		interm	GR_G	152	1.00
				snp		interm	GS_G	152	0.00
12	6,336,302			ins	1,984.0	interm	GR_E	138	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		AACACACACACACACACACACACACACACACAT	AACACACACACACACACACACACACACACACACAT	ins		interm	GS_E	138	0.01
				ins		slow	GR_F	138	1.00
				ins		slow	GS_F	138	0.00
15	1,850,364	CCACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACG	CCACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACG	ins	514.0	interm	GR_F	160	0.87
				ins		interm	GS_F	160	0.04
				ins		interm	GR_G	160	1.00
				ins		interm	GS_G	160	0.00
	2,498,470	GCACACACACACACACACACACACACACACACACACACAAC	GCACACACACACACACACACACACACACACACACACACACAAC	ins	384.0	interm	GR_F	76	1.00
				ins		interm	GS_F	160	0.00
				ins		slow	GR_E	160	0.93
				ins		slow	GS_E	160	0.00
16	272,187			ins	1,887.0	quick	GR_D	160	0.88

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGTGTGTGTGTGTGTGTGTGTTG	CGTGTGTGTGTGTGTGTGTGTTG	ins		quick	GS_D	160	0.00
				ins		interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.00
17	4,624,735	CGTGTGTGTGTGTGGTGC	CGTGTGTGTGTGTGGTGC	ins	282.0	quick	GR_G	97	1.00
				ins		quick	GS_G	47	0.44
				ins		interm	GR_F	160	1.00
				ins		interm	GS_F	111	0.38
01	217,144	TTGTGTGTGTGTGTGTGTGTGTGTGTTTGTGTGC	TTGTGTGTGTGTGTGTGTGTGTGTGTTTGTGTGC	ins	472.0	interm	GR_F	139	1.00
				ins		interm	GS_F	139	0.00
	258,451	TGGGGGGGGGGCGGGCGGAGG	TGGGGGGGGGGGGCGGGCGGAGG	ins	190.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	430,725	A	C	snp	1,623.0	quick	GR_F	143	1.00
				snp		quick	GS_F	143	0.00
	507,526	CTGTGTGTGTGTGTGTGTGTGTGTGTGTGC	CTGTGTGTGTGTGTGTGTGTGTGTGC	del	1,850.0	interm	GR_G	136	0.97
				del		interm	GS_G	136	0.01
	587,370	CGCAGCAGCAGCAGCAGCAGCAG	CGCAGCAGCAGCAGCAGCAG	ins	2,864.0	slow	GR_E	136	0.97
				ins		slow	GS_E	136	0.00
	754,962	ACCCCCCCCCTGG	ACCCCCCCCCCCTGG	ins	555.0	slow	GR_G	137	1.00
				ins		slow	GS_G	137	0.00
	1,006,293	ACACTCCACTCCACTCCACTCC	ACACTCCACTCCACTCCACTCCACTCC	ins	2,181.0	quick	GR_E	160	0.91
				ins		quick	GS_E	160	0.00
	1,282,990	C	T	snp	2,912.0	interm	GR_F	146	1.00
				snp		interm	GS_F	146	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	1,452,309	TTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTTA	TTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTTA	ins	1,004.0	slow	GR_F	142	1.00
				ins		slow	GS_F	142	0.00
	1,488,267	GCACACACAC ACACACACAC ACACAAG	GCACACACAC ACACACACAC ACACACAAG	ins	1,404.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	1,551,285	GCACACACAC ACACACACAC ACACACACAC ACACACAT	GCACACACAC ACACACACAC ACACACACAC ACACAT	del	1,217.0	slow	GR_G	137	1.00
				del		slow	GS_G	137	0.00
	2,276,397	G	A	snp	1,945.0	slow	GR_E	160	1.00
				snp		slow	GS_E	160	0.00
	3,226,598	CGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG	CGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG	ins	588.0	quick	GR_F	136	1.00
				ins		quick	GS_F	136	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGGTGGTGGTGGTGGTGGTGGGGAGT	GTGGTGGTGGTGGTGGTGGTGGTGGGAGT						
	3,631,914	GCACACACACACACG	GCACACACACACACG	ins	788.0	quick	GR_G	160	1.00
				ins		quick	GS_G	160	0.00
	3,944,117	T	C	snp	2,547.0	interm	GR_E	144	1.00
				snp		interm	GS_E	144	0.00
	4,699,130	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	ins	791.0	interm	GR_G	133	1.00
				ins		interm	GS_G	133	0.00
	4,829,965	GGTGTGTGTGTGTGTGTGTGTGTGTGT	GGTGTGTGTGTGTGTGTGTGTGTGTGTG	ins	217.0	interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGTGTGTGTG TGTGTGTGGC	TGTGTGTGTGT GGC						
	5,510,782	TGGGGGGGGG GGGTTAC	TGGGGGGGGG GGGGTTAC	ins	653.0	slow	GR_G	139	0.96
				ins		slow	GS_G	139	0.00
	5,636,548	C	A	snp	5,381.0	interm	GR_D	144	1.00
				snp		interm	GS_D	144	0.00
	6,196,146	CCCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACT	CCCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACCACACCAC ACT	del	403.0	quick	GR_F	160	1.00
				del		quick	GS_F	160	0.00
	7,199,539	CACAT	TACAC	complex	0.0	interm	GR_E	0	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				complex		interm	GS_E	134	0.04
	7,226,873	A	G	snp		interm	GR_E	0	1.00
				snp		interm	GS_E	133	0.08
	7,260,262	C		snp	5.0	interm	GR_G	5	1.00
				snp		interm	GS_G	160	0.31
	7,386,055	G	C	snp	1,898.0	interm	GR_D	160	1.00
				snp		interm	GS_D	160	0.00
	7,483,334	T	G	snp	2,469.0	interm	GR_E	141	1.00
				snp		interm	GS_E	141	0.00
	7,532,909	TACACACACACACACACACACACACACACACACACACAAACACACAT	TACACACACACACACACACACACACACACACACACACAACACACAT	ins	1,545.0	interm	GR_D	147	1.00
				ins		interm	GS_D	147	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	8,062,539	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	del	568.0	quick	GR_F	160	1.00
				del		quick	GS_F	160	0.03
02	344	TAAAACCCT	TAAACCCT	del	425.0	interm	GR_F	160	0.93
				del		interm	GS_F	160	0.00
	360	TAAACCC	TAAAACCC	ins	917.0	interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.02
	978,214	TCACACACACACACACACACACACACACACACACACACGCGCA	TCACACACACACACACACACACACACACACACACACACGCGCA	ins	1,688.0	slow	GR_G	160	0.97
				ins		slow	GS_G	160	0.00
	2,307,569			ins	527.0	interm	GR_F	139	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GGTGTGTGTGTGTGTGTGTGTGTGC	GGTGTGTGTGTGTGTGTGTGTGC	ins		interm	GS_F	139	0.00
	2,331,425	ATGTGTGTGTGTGTGTGTGTGTGTTTGTGTGTGC	ATGTGTGTGTGTGTGTGTGTGTGTTTGTGTGC	ins	1,393.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	2,658,571	AGGGGGGGGGGAGGTC	AGGGGGGGGGAGGTC	del	2,232.0	quick	GR_G	141	1.00
				del		quick	GS_G	141	0.00
	3,065,315	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	del	1,041.0	quick	GR_E	160	1.00
				del		quick	GS_E	160	0.00
	3,073,896	TTGTGTGTGTGTGTGTGTGTGTGT	TTGTGTGTGTGTGTGTGTGTGT	ins	1,234.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGTGTGTTTG TGA	GTGTGTGTGTT TGTGA						
	3,438,348	G	T	snp	2,188.0	interm	GR_E	160	1.00
				snp		interm	GS_E	160	0.00
	3,525,664	ACCCCCCCCC CGTTT	ACCCCCCCCC CCGTTT	ins	721.0	quick	GR_G	160	1.00
				ins		quick	GS_G	160	0.00
	3,762,754	GGTGTGTGTGT GTGTGTGTGTG TGC	GGTGTGTGTGT GTGTGTGTGTG TGTGC	ins	1,790.0	slow	GR_G	136	0.98
				ins		slow	GS_G	136	0.00
	4,728,172	TCTCCTCCTCC TCCTCCTCCTC CTCCTCCTCCT CCTCCTCCTCC A	TCTCCTCCTCC TCCTCCTCCTC CTCCTCCTCCT CCTCCTCCTCC TCCA	ins	527.0	quick	GR_E	160	1.00
				ins		quick	GS_E	160	0.00
	6,054,586	C	T	snp	3,762.0	slow	GR_F	160	1.00
				snp		slow	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	6,162,955	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	ins	474.0	slow	GR_G	148	0.95
				ins		slow	GS_G	148	0.00
	6,214,981	CGTGTGTGTGTGTGTGTGTGTGTTTGTGC	CGTGTGTGTGTGTGTGTGTGTTTGTGC	del	1,296.0	quick	GR_F	160	1.00
				del		quick	GS_F	160	0.00
	8,352,670	CCT	CCTAGGGCT	ins	2,464.0	quick	GR_G	140	0.94
				ins		quick	GS_G	140	0.00
	8,353,860	CCCCCAG	CCCGGGCCCCCAG	ins	3,078.0	interm	GR_E	160	1.00
				ins		interm	GS_E	0	
	8,409,942	GCCCCCCCCCCCCGAGA	GCCCCCCCCCCCCGAGA	ins	842.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
03	79,432	AATTTCA	GGTTTCCA	complex	0.0	quick	GR_F	0	1.00
				complex		quick	GS_F	139	0.00
	79,751	T	C	snp	693.0	quick	GR_F	135	1.00
				snp		quick	GS_F	135	0.01
	81,120			snp	0.0	quick	GR_F	0	1.00
				snp		quick	GS_F	160	0.08
	226,029	CGTGTGTGTGTGTGTGTGTGTGTTA	CGTGTGTGTGTGTGTGTGTGTTA	ins	1,150.0	interm	GR_E	160	1.00
				ins		interm	GS_E	160	0.00
	326,711	TGGGGGGGGGTTGGAA	TGGGGGGGGTTGGAA	del	1,166.0	quick	GR_G	160	1.00
				del		quick	GS_G	160	0.01
	470,045	TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	TCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	ins	1,644.0	quick	GR_G	139	0.91
				ins		quick	GS_G	139	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GCTGCTGCTGCTGCTGC	GCTGCTGCTGCTGCTGC						
	498,960	GAAAAGGT	GAAAGGT	del	1,070.0	slow	GR_F	142	1.00
				del		slow	GS_F	142	0.00
	514,072	TGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTTG	TGAGTGTGTGTGTGTGTGTGTGTGTGTGTTG	ins	1,160.0	quick	GR_F	160	0.98
				ins		quick	GS_F	160	0.01
	1,166,410	AAGTTGGCGG	AG	del	2,595.0	slow	GR_G	160	1.00
				del		slow	GS_G	160	0.00
	1,445,531	GACACACACACACACACACACACG	GACACACACACACACACACACACG	ins	939.0	interm	GR_E	136	0.97
				ins		interm	GS_E	136	0.00
	1,549,357			del	493.0	slow	GR_F	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GCACACACACACACACACACACAT	GCACACACACACACACACAT	del		slow	GS_F	160	0.00
	1,894,897	CGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTT	CGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTT	ins	451.0	quick	GR_E	160	1.00
				ins		quick	GS_E	160	0.00
	2,049,018	GCCACCCACCCACCCACCCACCCACCCACCCACGCACACG	GCCACCCACCCACCCACCCACCCACGCACACG	del	354.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.00
	2,116,424	GCCCCCCCCCCAGCGC	GCCCCCCCCCCCAGCGC	ins	650.0	interm	GR_G	144	1.00
				ins		interm	GS_G	144	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	2,664,628	GCACACACACACACACACACACACACACACACACACACG	GCACACACACACACACACACACACACACACACACACACG	ins	1,123.0	slow	GR_G	160	0.98
				ins		slow	GS_G	160	0.00
	3,446,876	GCCCCCCCCCCCCCATCAG	GCCCCCCCCCCCCCCATCAG	ins	2,098.0	slow	GR_E	137	0.99
				ins		slow	GS_E	137	0.02
	4,374,908	G	C	snp	3,512.0	slow	GR_F	160	1.00
				snp		slow	GS_F	160	0.00
	4,957,186		T	snp	2,816.0	slow	GR_E	160	0.98
				snp		slow	GS_E	160	0.00
	6,155,151			snp	1,598.0	interm	GR_F	141	0.94
				snp		interm	GS_F	141	0.00
	7,571,962			ins	1,105.0	slow	GR_E	145	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CCCACCACCACCACCACCACCACCACCACCACCACCACCAG	CCCACCACCACCACCACCACCACCACCACCACCACCACCAG	ins		slow	GS_E	145	0.00
	8,769,394	CGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	CGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	del	1,134.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.00
	9,159,155	ACACA	CCACC	complex	244.0	interm	GR_F	134	0.94
				complex		interm	GS_F	121	0.00
	9,185,690	T	C	snp	1,182.0	slow	GR_F	160	1.00
				snp		slow	GS_F	160	0.00
04	54,403			ins	525.0	interm	GR_E	160	0.96

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TTA	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTTA	ins		interm	GS_E	160	0.00
	236,253	TACAGGCACA GGCACAGGCA CAGGCACAGG CACAGGCACA GGCACAGGCA CAGGCACAGG CACAGGCA	TACAGGCACA GGCACAGGCA CAGGCACAGG CACAGGCACA GGCACAGGCA CAGGCA	del	878.0	quick	GR_D	160	0.90
				del		quick	GS_D	160	0.00
	1,182,986	CTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTTTC	CTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTTTC	del	399.0	interm	GR_F	160	0.94
				del		interm	GS_F	160	0.00
	1,612,570			ins	856.0	slow	GR_G	160	0.97

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TCCCCCCCCCCCTGCCTCTG	TCCCCCCCCCCCCGCCTCTG	ins		slow	GS_G	160	0.00
	1,624,635	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTGTGTGTGC	ins	1,836.0	interm	GR_D	151	1.00
				ins		interm	GS_D	151	0.00
	2,178,800	ACCCCCCCCCCCCTACG	ACCCCCCCCCCCCCTACG	ins	393.0	interm	GR_D	160	0.95
				ins		interm	GS_D	160	0.00
	2,366,026	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ATGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ins	899.0	interm	GR_D	160	1.00
				ins		interm	GS_D	160	0.00
	2,514,063	T	G	snp	2,872.0	interm	GR_F	153	1.00
				snp		interm	GS_F	153	0.00
	2,988,151	G	A	snp	2,851.0	quick	GR_D	139	0.88

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				snp		quick	GS_D	139	0.00
	3,024,945	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTC	TTGTGTGTGTGTGTGTGTGTGTGTGTGTC	del	617.0	slow	GR_G	141	0.96
				del		slow	GS_G	141	0.00
	3,251,425	C	A	snp	5,085.0	interm	GR_D	160	1.00
				snp		interm	GS_D	160	0.00
	230,918			snp	3,770.0	quick	GR_F	160	1.00
				snp		quick	GS_F	160	0.00
05	429,545	TGGGGGGGGGGTTGTTA	TGGGGGGGGGGGTTGTTA	ins	245.0	quick	GR_F	137	1.00
				ins		quick	GS_F	137	0.00
	908,228	ACAGCAGCAGCAGCAGCAGCAGA	ACAGCAGCAGCAGCAGCAGCAGA	ins	1,822.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	1,004,177			del	1,405.0	quick	GR_F	140	0.98

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		AGTGTGTGTGTGTGTGTGTGTGTTA	AGTGTGTGTGTGTGTGTGTTA	del		quick	GS_F	140	0.02
	1,137,275	TCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGGCA	TCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGGCA	ins	1,290.0	quick	GR_D	160	0.92
				ins		quick	GS_D	160	0.00
	2,130,722	AACACACACACACACACACACACG	AACACACACACACACACACACACG	ins	1,298.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
	2,297,133	ATGTGTGTGTGTGTGTGTGTGTGTGTGTTTGA	ATGTGTGTGTGTGTGTGTGTGTGTGTTTGA	ins	891.0	quick	GR_G	139	1.00
				ins		quick	GS_G	139	0.00
	2,357,694	CTGTGTGTGTGTGTGTGTGTGTGT	CTGTGTGTGTGTGTGTGTGTGT	ins	507.0	quick	GR_G	147	1.00
				ins		quick	GS_G	147	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGTGTGTGTG TTTGTGA	GTGTGTGTGTG TGTTTGTGA						
	2,620,554	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACG	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACG	ins	302.0	interm	GR_F	137	0.94
				ins		interm	GS_F	137	0.02
	2,735,633	G	T	snp	399.0	quick	GR_G	160	1.00
				snp		quick	GS_G	160	0.00
	2,756,792	CGGGGGGGGG GTAAA	CGGGGGGGGG TAAA	del	745.0	interm	GR_D	160	1.00
				del		interm	GS_D	160	0.07
	2,990,559	C	A	snp	3,187.0	slow	GR_E	137	1.00
				snp		slow	GS_E	137	0.02
06	307,123			ins	292.0	quick	GR_G	160	0.94

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	ins		quick	GS_G	160	0.00
	1,971,272	GCACACACACACACACACACACACACACACACACAT	GCACACACACACACACACACACACACACACACACAT	ins	182.0	interm	GR_E	137	1.00
				ins		interm	GS_E	137	0.00
	2,462,667	AACACACACACACACACACACACACACACAT	AACACACACACACACACACACACACACACAT	del	2,552.0	quick	GR_E	143	1.00
				del		quick	GS_E	143	0.00
	3,530,061	TGGGGGGGGGGCACT	TGGGGGGGGGGGCACT	ins	986.0	slow	GR_F	136	1.00
				ins		slow	GS_F	136	0.00
	4,075,824			del	639.0	interm	GR_D	139	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGC	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGC	del		interm	GS_D	139	0.00
	4,599,211	CTAT	CT	del	4,510.0	quick	GR_E	137	1.00
				del		quick	GS_E	137	0.00
	5,097,107	ACCCCCCCCC ACCT	ACCCCCCCCC CACCT	ins	1,685.0	slow	GR_G	160	0.98
				ins		slow	GS_G	160	0.00
	5,480,613	ATGTGTGTGTG TGTGTGTGTGA	ATGTGTGTGTG TGTGTGTGTGT GTGA	ins	1,497.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	5,711,460	GACACACACA CACACACACA CACACACACA CACACACACA CACACACAT	GACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CAT	ins	388.0	quick	GR_F	134	1.00
				ins		quick	GS_F	134	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	6,375,099	GAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGG	GAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGG	del	364.0	quick	GR_G	136	1.00
				del		quick	GS_G	136	0.00
	7,318,185	GACACACACACACACACACACACACACACACACACACAG	GACACACACACACACACACACACACACACACACACACACACACACACACAG	ins	1,996.0	interm	GR_D	160	0.97
				ins		interm	GS_D	160	0.00
	7,529,017	GCGGTGGGGCC	GC	del	174.0	quick	GR_F	153	1.00
				del		quick	GS_F	153	0.00
	7,633,103	GCACACACACACACACACG	GCACACACACACACACACACG	ins	3,063.0	quick	GR_E	142	1.00
				ins		quick	GS_E	142	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	8,237,852	ATGTGTGTGTGTGTGTGTGTGTGTGTGC	ATGTGTGTGTGTGTGTGTGTGTGTGC	del	800.0	interm	GR_F	160	1.00
				del		interm	GS_F	160	0.02
	8,629,983	CATACACACACACACACACG	CATACACACACACACACACG	ins	2,659.0	slow	GR_F	138	1.00
				ins		slow	GS_F	138	0.00
07	21,924	TGC	TC	del	0.0	quick	GR_F	0	1.00
				del		quick	GS_F	160	0.03
	24,590	G	T	snp		quick	GR_F	0	1.00
				snp		quick	GS_F	160	0.04
	25,017	A	C	snp		quick	GR_F	0	1.00
				snp		quick	GS_F	160	0.09
	189,532	GGTGTGTGTGTGTGTGTGTGTGTGTG	GGTGTGTGTGTGTGTGTGTGTGTG	ins	1,006.0	interm	GR_F	160	0.95
				ins		interm	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TGTGTGTGTGTGTGTTTA	TGTGTGTGTGTGTGTTTA						
	359,246	ACCCCCCCCCCCGCGCACA	ACCCCCCCCCCGCGCACA	del	2,380.0	slow	GR_G	160	1.00
				del		slow	GS_G	160	0.00
	640,757	ATGTGTGTGTGTGTGTGTGTGTGTGGGTGTGTGTTT	ATGTGTGTGTGTGTGTGTGTGGGTGTGTGTTT	del	558.0	slow	GR_G	160	1.00
				del		slow	GS_G	160	0.00
	1,399,520	C	T	snp	3,556.0	interm	GR_D	142	1.00
				snp		interm	GS_D	142	0.00
	1,994,476	AACACACACACACACACACACACAT	AACACACACACACACACACAT	del	1,617.0	quick	GR_G	160	1.00
				del		quick	GS_G	160	0.00
	2,130,886	AAG	AAGCAG	ins	2,528.0	interm	GR_F	133	0.93
				ins		interm	GS_F	133	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	2,234,373	AACACACACACACACACACACACACG	AACACACACACACACACG	del	2,376.0	slow	GR_G	160	1.00
				del		slow	GS_G	160	0.01
	2,307,693	TGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAC	TGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAGCACCAC	del	1,185.0	slow	GR_G	139	1.00
				del		slow	GS_G	139	0.00
	2,412,917	GCACACACACACACACACACACATAC	GCACACACACACACACACACATAC	ins	1,634.0	quick	GR_F	134	0.98
				ins		quick	GS_F	134	0.00
	2,674,272	TTGTGTGTGTGTGTGTGTGTGTGTA	TTGTGTGTGTGTGTGTGTGTA	ins	2,627.0	interm	GR_G	137	0.99
				ins		interm	GS_G	137	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	2,960,519	T	G	snp	1,888.0	quick	GR_E	142	0.91
				snp		quick	GS_E	142	0.02
	3,324,261	A		snp	2,627.0	interm	GR_F	160	1.00
				snp		interm	GS_F	160	0.00
	3,955,178	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTA	ins	574.0	quick	GR_G	142	1.00
				ins		quick	GS_G	142	0.00
	4,174,874	GCACACACACACACACACACAT	GCACACACACACACACACAT	ins	1,440.0	slow	GR_E	137	1.00
				ins		slow	GS_E	137	0.00
	4,783,188	AAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCA	AAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCA	del	292.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCG	GGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCG						
	4,848,576	AACACACACACACACACACACACG	AACACACACACACACACACACACG	ins	609.0	interm	GR_E	160	0.96
				ins		interm	GS_E	160	0.00
	5,054,903	GCACACACACACACACACACACG	GCACACACACACACACACACACG	ins	832.0	slow	GR_F	160	0.97
				ins		slow	GS_F	160	0.00
	5,496,794	GGGGGCAGGGGCAGGGGCAGGGGCAGGGGCAG	GGGGGCAGGGGCAGGGGCAGGGGCAGGGGCAG	ins	503.0	slow	GR_G	134	0.86
				ins		slow	GS_G	134	0.02
	5,975,301	A	G	snp	2,963.0	interm	GR_G	160	1.00
				snp		interm	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	6,095,652	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGC	GTGTGTGTGTGTGTGTGTGTGTGTGTGC	snp	256.0	slow	GR_F	116	0.93
				snp		slow	GS_F	160	0.04
	6,251,047	CGTGTGTGTGTGTGTGTGTGTGTGTGTA	CGTGTGTGTGTGTGTGTGTGTGTGTGTA	ins	829.0	quick	GR_E	138	1.00
				ins		quick	GS_E	138	0.00
	6,254,608	TCACACACACACACACACACACACAG	TCACACACACACACACACACACACAG	ins	455.0	quick	GR_G	160	0.96
				ins		quick	GS_G	160	0.00
08	416,680	TCACACACACACACACACACACACACACACACACACACACACCT	TCACACACACACACACACACACACACACACACACACACACACCT	ins	417.0	interm	GR_E	142	1.00
				ins		interm	GS_E	142	0.00
	496,202			ins	778.0	slow	GR_F	160	0.97

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	ins		slow	GS_F	160	0.00
	594,711	G	C	snp	1,459.0	interm	GR_E	153	0.95
				snp		interm	GS_E	153	0.00
	720,914	ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGG	ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGG	del	1,010.0	quick	GR_G	160	1.00
				del		quick	GS_G	160	0.00
	1,334,262	AGGGGGGGGGGGTCGTT	AGGGGGGGGGGTCGTT	del	2,434.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.02
	1,958,906			ins	541.0	slow	GR_G	160	0.95

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT	ins		slow	GS_G	160	0.00
	2,940,149	T	A	snp	1,613.0	interm	GR_E	147	1.00
				snp		interm	GS_E	147	0.01
	3,214,279	GACACACACACACACACACACCT	GACACACACACACACACACACACCT	ins	2,850.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
	3,421,934	GCACACACACACACACACACACACACACACACACACACACACACACACACACACACAG	GCACACACACACACACACACACACACACACACACACACACACACAC	ins	214.0	quick	GR_G	160	0.93
				ins		quick	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
			ACACACACACAG						
	3,460,230	CCACACACACACACACACACACACACACACACACACACACACACAG	CCACACACACACACACACACACACACACACACACACACACACAG	ins	1,051.0	interm	GR_G	135	0.98
				ins		interm	GS_G	135	0.00
	3,557,195	AGGGGGGGGGGGTAA	AGGGGGGGGGGGGTAA	ins	224.0	interm	GR_F	146	0.92
				ins		interm	GS_F	146	0.04
	3,948,362	AC	AGACGGA	complex	3,805.0	quick	GR_E	133	1.00
				complex		quick	GS_E	133	0.00
	4,031,353	TTGTGTGTGTGTGTGTGTGTGTGTGTGT	TTGTGTGTGTGTGTGTGTGTGTGTGCGT	ins	880.0	interm	GR_G	134	0.97
				ins		interm	GS_G	134	0.00
	4,197,249	GCACACACACACACACACACACACACACACACACACACACACAC	GCACACACACACACACACACACACACACACACACACACACACAC	ins	437.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ACACACACACACACACACACG	ACACACACACACACACACACACG						
	4,397,893	GGTGTGTGTGTGTGTGTGTGTGTA	GGTGTGTGTGTGTGTGTGTGTA	ins	1,574.0	interm	GR_F	160	0.98
				ins		interm	GS_F	160	0.00
09	1,209,544	GCCTG	GCCCTG	ins	2,183.0	interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.00
	1,216,520	G	C	snp	0.0	interm	GR_F	0	1.00
				snp		interm	GS_F	160	0.00
	1,417,560	TCCTCCACCTCCACCTCCACCTCCACCTCCACCCCACCTCCACCTCCACCTCCACCTCCACCTCCAT	TCCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCATCCTCCAT	ins	350.0	slow	GR_F	144	1.00
				ins		slow	GS_F	144	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	1,579,392	TGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGA	TGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGA	del	663.0	slow	GR_E	140	1.00
				del		slow	GS_E	140	0.00
	1,710,470	ATGTGTGTGTGTGTGTGTGTGTGTA	ATGTGTGTGTGTGTGTGTGTGTA	ins	638.0	interm	GR_E	148	0.94
				ins		interm	GS_E	148	0.00
	3,816,199	ACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTTCTTCCCTTCTCG	ACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTTCTTCTCCTTCTCG	ins	358.0	quick	GR_G	160	0.90
				ins		quick	GS_G	160	0.00
	3,856,765	CTGTGTGTGTGTGTGTGTGTGTGT	CTGTGTGTGTGTGTGTGTGTGT	ins	974.0	quick	GR_G	160	0.97
				ins		quick	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGTGTGTGTG TGTC	GTGTGTGTGTG TGTGTC						
	4,002,402	C	T	snp	2,664.0	quick	GR_G	160	1.00
				snp		quick	GS_G	160	0.00
	4,648,989		G	snp	251.0	interm	GR_E	160	0.88
				snp		interm	GS_E	160	0.40
	4,648,994			snp	129.0	interm	GR_E	126	0.92
				snp		interm	GS_E	129	0.38
	4,688,496	GCACACACAC ACACACACAC AT	GCACACACAC ACACACACAC ACAT	ins	1,600.0	quick	GR_F	160	1.00
				ins		quick	GS_F	160	0.00
	4,854,741	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG	ins	742.0	slow	GR_G	138	1.00
				ins		slow	GS_G	138	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TGTGTGTGTGT GTGTGTTA	TGTGTGTGTGT GTGTGTGTTA						
	5,411,399	TGGGGGGGGG GACA	TGGGGGGGGG GGACA	ins	2,440.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
	5,413,779	ACCCCCCCCC CTGCA	ACCCCCCCCC CCTGCA	ins	1,019.0	quick	GR_D	138	1.00
				ins		quick	GS_D	138	0.00
	5,418,408	GGTGTGTGTGT GTGTGTGTGTG TGTGTTG	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTTG	ins	564.0	quick	GR_E	160	1.00
				ins		quick	GS_E	160	0.00
	5,524,304	GCCCCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACG	ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACACC ACACCACG	complex	396.0	slow	GR_G	133	1.00
				complex		slow	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	5,569,262	GCTGCCGCCGCGCGTGCGGCCCAGCTGGCTCGGGGGCC	GC	del	672.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.00
	5,712,694	CGTGTGTGTGTGTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGC	del	2,104.0	interm	GR_D	147	1.00
				del		interm	GS_D	147	0.00
	5,926,830	ACCCCCCCCCCCCATTG	ACCCCCCCCCCCATTG	del	565.0	interm	GR_G	153	1.00
				del		interm	GS_G	153	0.05
	6,143,181	GGCCGGACG	GG	del	407.0	interm	GR_F	160	0.91
				del		interm	GS_F	160	0.00
	6,210,316	C	T	snp	3,564.0	interm	GR_G	160	1.00
				snp		interm	GS_G	160	0.00
	6,426,674	CTGCACGAGG	CTGCACGAGTGCACGAGG	ins	1,836.0	quick	GR_F	160	1.00
				ins		quick	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
10	388,358	ACCCCCCCCCCCACCTCACAG	ACCCCCCCCCCCCACCTCACAG	ins	1,621.0	quick	GR_D	143	0.93
				ins		quick	GS_D	143	0.00
	660,244	C	T	snp	3,915.0	quick	GR_D	160	0.95
				snp		quick	GS_D	160	0.00
	965,483	G		snp	4,390.0	interm	GR_G	145	1.00
				snp		interm	GS_G	145	0.00
	1,564,912	GC	GTAT	complex	2,730.0	quick	GR_E	149	1.00
				complex		quick	GS_E	149	0.00
	1,777,785	TCACACACACACACACACACACACACACACACACACACACACACACACACACACACG	TCACACACACACACACACACACACACACACACACACACACACACACACACACACACACG	ins	189.0	quick	GR_F	86	0.92
				ins		quick	GS_F	160	0.00
	1,938,590			ins	821.0	interm	GR_G	139	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TACATAGCTTG CACT	TACATAGCTTG CACACATAGCT TGCACT	ins		interm	GS_G	139	0.00
	2,113,406	GCCCCCCCCC TCCTCCTGCT	GCCCCCCCCC CTCCTCCTGCT	ins	1,266.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	2,534,998	TCGCCGACCG GCTCGC	TC	del	2,449.0	quick	GR_D	142	0.92
				del		quick	GS_D	142	0.00
	2,932,776	TACACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACACAC ACACCT	TACACACACAC ACACACACAC ACACACACAC ACACACACAC ACACACACAC ACCT	del	372.0	interm	GR_E	160	0.89
				del		interm	GS_E	160	0.03
	3,226,288	AGGGGCA	AGGCA	del	3,745.0	slow	GR_G	143	1.00
				del		slow	GS_G	143	0.00
	3,285,917			ins	1,603.0	slow	GR_F	137	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ATGTGTGTGTGTGTGTGTGTGTGTTTGTA	ATGTGTGTGTGTGTGTGTGTGTTTGTA	ins		slow	GS_F	137	0.00
	4,329,337	ACCCACACACACACACACACACACG	ACCCACACACACACACACACACACACG	ins	1,254.0	slow	GR_F	145	1.00
				ins		slow	GS_F	145	0.00
	4,425,835	C	A	snp	2,454.0	slow	GR_E	160	1.00
				snp		slow	GS_E	160	0.00
	4,731,857	CCACACACACACACACACACACACACACACACG	CCACACACACACACACACACACACACACG	del	1,493.0	interm	GR_G	160	1.00
				del		interm	GS_G	160	0.01
	4,840,394	AACACACACACACACACG	AACACACACACACACACACG	ins	1,355.0	interm	GR_D	160	0.98
				ins		interm	GS_D	160	0.00
	5,047,265			ins	824.0	interm	GR_G	137	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GCACACACACACACACACACACACACACACT	GCACACACACACACACACACACACACACACACT	ins		interm	GS_G	137	0.02
	5,246,752	G	A	snp	2,862.0	interm	GR_F	160	1.00
				snp		interm	GS_F	160	0.00
	5,647,470	TTGCTGCTGCTGCTGCTGCTGCTGCTGCG	TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG	ins	2,733.0	interm	GR_D	160	1.00
				ins		interm	GS_D	160	0.00
	5,660,729	C	A	snp	3,199.0	interm	GR_D	146	1.00
				snp		interm	GS_D	146	0.00
	5,831,355	G	C	snp	2,490.0	interm	GR_D	160	1.00
				snp		interm	GS_D	160	0.00
	6,137,866			del	1,811.0	slow	GR_F	160	0.98

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	GTGCTGCTGCTGCTGCTGCTGCTGCT	del		slow	GS_F	160	0.00
	6,401,976	ACCCCCCCCCCCCCGCAC	ACCCCCCCCCCCCCGCAC	ins	590.0	interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.00
	6,475,403	TTGTGTGTGTGTGTGTGTGTGTGTTTGGA	TTGTGTGTGTGTGTGTGTGTGTTTGGA	ins	1,138.0	quick	GR_G	143	1.00
				ins		quick	GS_G	143	0.00
	6,719,206	TAG	TG	del	1,344.0	interm	GR_E	160	0.92
				del		interm	GS_E	160	0.00
11	950	T	C	snp	111.0	slow	GR_G	111	0.92
				snp		slow	GS_G	61	0.00
	965			snp	118.0	slow	GR_G	118	0.92
				snp		slow	GS_G	53	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	972	CAT	CT	del	123.0	slow	GR_G	124	0.91
				del		slow	GS_G	86	0.00
	984	A	C	snp	184.0	slow	GR_G	146	0.92
				snp		slow	GS_G	86	0.00
	1,008	AAA	CAG	complex	219.0	slow	GR_G	147	1.00
				complex		slow	GS_G	31	0.00
	1,033	C	G	snp	214.0	slow	GR_G	160	1.00
				snp		slow	GS_G	63	0.00
	1,296,346	TACACACACACACACACACACACACACACACACACACACACACACACACACACACG	TACACACACACACACACACACACACACACACACACACACACACACACACACG	del	469.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.03

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	1,473,139	ATGTGTGTGTGTGTGTGTGTTTGC	ATGTGTGTGTGTGTGTGTGTTTGC	ins	838.0	interm	GR_E	160	1.00
				ins		interm	GS_E	160	0.00
	1,813,533	G	T	snp	3,686.0	quick	GR_D	148	0.95
				snp		quick	GS_D	148	0.00
	2,148,461	TCACACACACACACACACACACACACACACACACACACACACAT	TCACACACACACACACACACACACACACACACACACACACACAT	ins	644.0	interm	GR_E	135	0.93
				ins		interm	GS_E	135	0.00
	2,436,547	AACACACACACACACACACACACACACACACACACACACACACACACACCCG	AACACACACACACACACACACACACACACACACACACACACACACACCCG	del	251.0	slow	GR_F	160	1.00
				del		slow	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	3,290,360	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTA	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTA	del	603.0	slow	GR_G	160	1.00
				del		slow	GS_G	160	0.00
	3,304,362	CCACACACAC ACACACACAC ACACACACAC ACACAAACAC ACAAAAACAC ACACACACAC ACACACACAC ACACG	CCACACACAC ACACACACAC ACACACACAC ACAAACACAC AAAAACACAC ACACACACAC ACACACACAC ACG	del	397.0	quick	GR_G	160	1.00
				del		quick	GS_G	160	0.00
	3,365,906	GCACACACAC ACACACACAC ACACACACAC ACACACACG	GCACACACAC ACACACACAC ACACACACAC ACACACACAC G	ins	311.0	interm	GR_E	138	1.00
				ins		interm	GS_E	138	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	4,229,438	CGTGTGTGTGT GTGTGTGTGTG TTTGTGC	CGTGTGTGTGT GTGTGTGTGTG TGTTTGTGC	ins	1,272.0	quick	GR_F	160	0.96
				ins		quick	GS_F	160	0.01
	4,299,599	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTTT GTGTGTGTGTG TGTGTGTGTGT TA	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT TTGTGTGTGTG TGTGTGTGTGT GTTA	ins	327.0	interm	GR_G	160	0.94
				ins		interm	GS_G	160	0.02
12	137,918	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACAT	AACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACAT	del	414.0	interm	GR_G	160	1.00
				del		interm	GS_G	160	0.00
	702,175			ins	1,573.0	slow	GR_F	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC	ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC	ins		slow	GS_F	160	0.00
	829,202	C	T	snp	581.0	slow	GR_G	160	0.88
				snp		slow	GS_G	160	0.03
	829,232			snp	925.0	slow	GR_G	160	0.88
				snp		slow	GS_G	160	0.00
	1,111,229	ACGCACACACACACACACACACACACACACACACACACACACACACACT	ACGCACACACACACACACACACACACACACACACACACACACACACACACT	ins	900.0	interm	GR_F	135	1.00
				ins		interm	GS_F	135	0.06
	1,119,261			ins	2,242.0	slow	GR_E	138	0.97

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TACACACACACTACACACACAC ACACACACAC ACACACACAC AAT	TACACACACAC ACACACACAC ACACACACAC ACAAT	ins		slow	GS_E	138	0.00
	1,168,724	AGTGTGTGTGT GTGTGTGTGTG TGTGTGTA	AGTGTGTGTGT GTGTGTGTGTG TGTGTGTA	ins	371.0	quick	GR_F	144	1.00
				ins		quick	GS_F	144	0.00
	1,653,521	GACACACACA CACACACACA CACACACACA CACACACACC G	GACACACACA CACACACACA CACACACACA CACACACACA CCG	ins	985.0	slow	GR_E	160	0.95
				ins		slow	GS_E	160	0.00
	1,653,565	GCACACACAC ACACACACAC ACACACACAC ACACACACAC CG	GCACACACAC ACACACACAC ACACACACAC ACACACACCG	del	529.0	slow	GR_F	148	1.00
				del		slow	GS_F	148	0.00
	2,748,952			ins	355.0	quick	GR_G	150	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		AGTGTGTGTGTGTGTGTGTGTGTGTGTGTC	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTC	ins		quick	GS_G	150	0.00
	2,999,164	CAAACACACACACACACACACACACACACACACACAAG	CAAACACACACACACACACACACACACACACACAAG	ins	916.0	interm	GR_E	160	1.00
				ins		interm	GS_E	160	0.00
	3,272,417	AGTGTGTGTGTGTGTGTA	AGTGTGTGTGTGTGTGTA	ins	2,775.0	quick	GR_E	139	1.00
				ins		quick	GS_E	139	0.00
	3,768,419	C	T	snp	4,215.0	interm	GR_E	146	1.00
				snp		interm	GS_E	146	0.00
	4,370,490	TGGGGGGGGGCGTTA	TGGGGGGGGGGCGTTA	ins	1,096.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
	4,618,179			ins	827.0	interm	GR_F	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGGGGGGGGGGGGGTTAC	CGGGGGGGGGGGGGTTAC	ins		interm	GS_F	160	0.00
	6,222,214	A	C	snp	3,226.0	interm	GR_F	141	1.00
				snp		interm	GS_F	141	0.00
	6,328,210	TCCCGCCCGCCCGCCCGCCCGCCCTCGCA	TCCCGCCCGCCCGCCCGCCCGCCCCTCGCA	ins	484.0	quick	GR_F	137	1.00
				ins		quick	GS_F	137	0.00
	6,575,282	GAGCGTAGCGTT	GAGCGTT	del	3,958.0	quick	GR_F	136	1.00
				del		quick	GS_F	136	0.00
	9,088,322	TGGGGGGGGGGGGGTTGTAA	TGGGGGGGGGGGGGTTGTAA	ins	892.0	interm	GR_F	135	0.97
				ins		interm	GS_F	135	0.00
13	233,788	CGTGTGTGTGTGTGTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGTG	ins	1,167.0	interm	GR_G	141	1.00
				ins		interm	GS_G	141	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
			TGTGTGTGTGT GC						
	264,383	TACACACACAC ACACACACAC ACACACACAC ACACACCT	TACACACACAC ACACACACAC ACACACACAC ACACACACCT	ins	1,563.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	518,989	TACACACACAC ACACACACAC ACACACACAC G	TACACACACAC ACACACACAC ACACACACAC ACG	ins	1,120.0	interm	GR_G	138	1.00
				ins		interm	GS_G	138	0.00
	811,394	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGA	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGA	ins	180.0	interm	GR_E	135	0.93
				ins		interm	GS_E	135	0.00
	2,354,166			ins	654.0	interm	GR_D	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ACCCCCCCCACACACACACACACACG	ACCCCCCCCCCACACACACACACACG	ins		interm	GS_D	160	0.00
	2,576,026	GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTC	GGTGTGTGTGTGTGTGTGTGTGTGTGTGTC	ins	977.0	interm	GR_D	160	0.95
				ins		interm	GS_D	160	0.00
	3,088,622	GCCCCCCCCCCCCACTC	GCCCCCCCCCCCACTC	del	1,669.0	interm	GR_G	133	0.98
				del		interm	GS_G	133	0.02
	3,343,294	G	C	snp	4,651.0	interm	GR_D	160	0.99
				snp		interm	GS_D	160	0.00
	3,768,290	CCACACACACACACACACACACACACACACAG	CCACACACACACACACACACACACACACAG	del	1,258.0	slow	GR_F	137	1.00
				del		slow	GS_F	137	0.00
	3,851,184	CCGGC	CCGGCGGC	ins	2,757.0	interm	GR_F	142	0.99

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				ins		interm	GS_F	142	0.00
	4,094,783	TGCCGCCGCCGCCGCCGCCGCT	TGCCGCCGCCGCCGCCGCT	del	1,012.0	slow	GR_F	138	1.00
				del		slow	GS_F	138	0.00
	4,638,826	G	A	snp	3,532.0	slow	GR_G	160	1.00
				snp		slow	GS_G	160	0.00
	4,693,224	GCACACACACACACACACACACACAT	GCACACACACACACACACACACACACAT	ins	833.0	interm	GR_G	142	1.00
				ins		interm	GS_G	142	0.00
	4,794,827	AACACACACACACACACACACACACACACG	AACACACACACACACACACACACACACACAG	ins	2,215.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	5,163,079	AACACACACACACACACACACACACACA	AACACACACACACACACACACACACACA	ins	979.0	interm	GR_E	160	0.97
				ins		interm	GS_E	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CACACACACACACG	CACACACACACACG						
	5,177,388	GCACACACACACACACACACACACACACACACACACACACG	GCACACACACACACACACACACACACACACACACACACG	ins	1,048.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	5,177,487	AACACACACACACACACACACACACACACG	AACACACACACACACACACACG	del	2,053.0	interm	GR_G	160	1.00
				del		interm	GS_G	160	0.01
14	208	TAAAAAAAACCCAAA	AAAAAAG	complex	361.0	slow	GR_G	145	1.00
				complex		slow	GS_G		
	311	GAAAA	AAAAC	complex	1,320.0	slow	GR_G	160	0.98
				complex		slow	GS_G	160	0.12
	330	ACCTA	ACCCAAA	complex	1,451.0	slow	GR_G	160	0.97
				complex		slow	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	227,235	GCACACACAC ACACACACAC ACACACACAC ACACACAT	GCACACACAC ACACACACAC ACACACACAC ACACACACAT	ins	402.0	quick	GR_F	144	1.00
				ins		quick	GS_F	144	0.00
	1,448,634	CGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGA	CGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG GTGGTGGTGG TGGTGGTGGA	ins	505.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	1,603,279	CCACACACAC ACACACACAC ACACACACAC ACACACACAA C	CCACACACAC ACACACACAC ACACACACAC ACACACACAC AAC	ins	539.0	interm	GR_D	160	1.00
				ins		interm	GS_D	160	0.00
	1,651,829	GGCCAGCGCC AGCGCCAGCG CCAGCGCCAG	GGCCAGCGCC AGCGCCAGCG CCAGCGCCAG	del		quick	GR_F	154	1.00
				del		quick	GS_F	154	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CGCCAGCGCC AGCGCCAGCG CCAGCGCCAG CGCCAGCGCC AGCGCCAGCG CCAGCGCCAG CGCCAGCGCC AGCGCCAGCG CCAGCG	CGCCAGCGCC AGCGCCAGCG CCAGCGCCAG CGCCAGCGCC AGCGCCAGCG CCAGCGCCAG CGCCAGCGCC AGCGCCAGCG						
	1,681,521	TACACACACACT ACAAACAT	TACACACACAC ACACAAACAT	ins	2,383.0	interm	GR_G	138	1.00
				ins		interm	GS_G	138	0.00
	1,794,162	TCACACACACAT CACACACACA CACACACACA CACACAT	TCACACACACA CACACACACA CACACACACA CACACACAT	ins	1,217.0	slow	GR_G	160	0.98
				ins		slow	GS_G	160	0.00
	2,510,754	A	C	snp	1,427.0	interm	GR_G	142	1.00
				snp		interm	GS_G	142	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	2,917,901	GCACACACACACACACACG	GCACACACACACACACACACG	ins	1,979.0	slow	GR_F	137	1.00
				ins		slow	GS_F	137	0.00
	2,925,005	TGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGT	TGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGT	ins	724.0	interm	GR_G	160	0.97
				ins		interm	GS_G	160	0.00
	3,218,703	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTG	ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTG	del	797.0	quick	GR_G	160	1.00
				del		quick	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	3,780,281	GCACACACACACACACG	GCACACACACACACACG	ins	2,000.0	quick	GR_F	160	0.99
				ins		quick	GS_F	160	0.01
15	298,175	GCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTG	GCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTG	ins	1,046.0	slow	GR_G	137	1.00
				ins		slow	GS_G	137	0.00
	385,765	AACACACACACACACACACACACACACACACACACACACACACACACACG	AACACACACACACACACACACACACACACACACACACACACACACACACG	del	597.0	interm	GR_G	160	1.00
				del		interm	GS_G	160	0.00
	434,672	CTGTGTGTGTGTGTGTGTGC	CTGTGTGTGTGTGTGTGTGC	ins	1,229.0	interm	GR_D	136	0.96
				ins		interm	GS_D	136	0.00
	555,528	CGTGTGTGTGTGTGTGTGTGTGTGTGTGT	CGTGTGTGTGTGTGTGTGTGTGTGTGTGT	ins	211.0	interm	GR_E	135	1.00
				ins		interm	GS_E	135	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GTGTGTGTGTGTGTGTGTGTTTA	GTGTGTGTGTGTGTGTGTGTTTTA						
	602,821	ATGTGTGTGTGTGTGTGTGTGTGTGTGTTGTG	ATGTGTGTGTGTGTGTGTGTGTGTTTGTG	ins	1,016.0	interm	GR_F	160	1.00
				ins		interm	GS_F	160	0.00
	1,204,148	CTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGA	CTGTGTGTGTGTGTGTGTGTGTGTGTTTGA	ins	875.0	slow	GR_G	160	0.97
				ins		slow	GS_G	160	0.00
	1,341,883	A	G	snp	3,846.0	quick	GR_E	143	1.00
				snp		quick	GS_E	143	0.00
	1,499,813		C	snp	949.0	quick	GR_G	145	1.00
				snp		quick	GS_G	145	0.02
	1,547,075	C	G	snp	3,091.0	slow	GR_G	160	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
				snp		slow	GS_G	160	0.00
	1,794,425	TACACACACACACACACACAG	TACACACACACACACACAG	ins	1,421.0	interm	GR_G	160	0.96
				ins		interm	GS_G	160	0.00
	2,583,304	CACAGACAGACAGACAGACAGA	CACAGACAGACAGACAGACAGA	ins	1,944.0	slow	GR_E	145	0.94
				ins		slow	GS_E	145	0.00
	2,954,314	G	A	snp	421.0	slow	GR_F	160	1.00
				snp		slow	GS_F	160	0.00
	2,954,330	CTGA	ATGG	complex	710.0	slow	GR_F	149	1.00
				complex		slow	GS_F	149	0.00
	3,496,890	GAATAATAATAATAATAATAATAATAATAG	GAATAATAATAATAATAATAATAATAATAG	ins	2,338.0	interm	GR_F	130	0.99
				ins		interm	GS_F	130	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	3,796,093	T	A	snp	129.0	interm	GR_F	128	1.00
				snp		interm	GS_F	114	0.22
16	437,347	C	T	snp	3,879.0	slow	GR_F	160	1.00
				snp		slow	GS_F	160	0.00
	475,776	AACACACACA CACACACG	AACACACACA CACACACACG	ins	1,455.0	slow	GR_G	160	1.00
				ins		slow	GS_G	160	0.00
	493,316	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTTA	GGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTTA	ins	1,119.0	slow	GR_E	142	1.00
				ins		slow	GS_E	142	0.23
	519,168	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTC	CGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTC	ins	644.0	interm	GR_E	136	1.00
				ins		interm	GS_E	136	0.00
	777,925			del	362.0	interm	GR_E	136	1.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	ATGTGTGTGTGTGTGTGTGTGTGTA	del		interm	GS_E	136	0.00
	797,293	TGGGGGGGGGGCTTT	TGGGGGGGGGCTTT	del	1,846.0	slow	GR_F	160	0.97
				del		slow	GS_F	160	0.00
	801,084	TGCCCCGCCCCGCCCT	TGCCCT	del	1,559.0	interm	GR_E	160	1.00
				del		interm	GS_E	160	0.00
	1,450,311	TACACACACACTACACG	TACACACACACACG	ins	1,456.0	interm	GR_D	144	1.00
				ins		interm	GS_D	144	0.00
	2,044,962	AGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCA	AGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCAGGCA	del	2,099.0	slow	GR_E	160	1.00
				del		slow	GS_E	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GGCAGGCAGGCAGGCAGGCAGGCAGGCAGG	GGCAGGCAGGCAGGCAGGCAGGCAGGCAGG						
	3,239,131	GGGGGGGAAGC	GGGGGGGGGCGGGGGGAAGC	ins	243.0	slow	GR_E	38	1.00
				ins		slow	GS_E	0	
	3,813,463	GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	del	256.0	slow	GR_G	136	0.88
				del		slow	GS_G	136	0.00
	4,034,209	GCCGGC	GCCCGTG	complex	2,114.0	interm	GR_E	141	1.00
				complex		interm	GS_E	141	0.00
	4,382,782	G	A	snp	3,819.0	slow	GR_G	160	1.00
				snp		slow	GS_G	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	4,843,947	TGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	TGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	ins	457.0	interm	GR_E	160	0.85
				ins		interm	GS_E	160	0.00
	5,078,695	TTGTGTGTGTGTGTGTGTATG	TTGTGTGTGTGTGTGTGTGTATG	ins	2,074.0	slow	GR_F	137	1.00
				ins		slow	GS_F	137	0.00
	5,844,576	T	C	snp	4,211.0	interm	GR_G	160	1.00
				snp		interm	GS_G	160	0.00
	6,418,147	GCACACACACACACACACACAG	GCACACACACACACACACACAG	ins	689.0	slow	GR_G	142	1.00
				ins		slow	GS_G	142	0.00
	6,726,301	CGTGTGTGTGTGTGTGTGTGTGC	CGTGTGTGTGTGTGTGTGTGTGC	ins	1,731.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	6,991,246			ins	1,229.0	interm	GR_G	143	0.94

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		AACACACACACACACACACACACACACACACACACACACACACG	AACACACACACACACACACACACACACACACACACACACACG	ins		interm	GS_G	143	0.00
	7,008,482	CGCTGCTGCTGCTGCTGCTGCTG	CGCTGCTGCTGCTGCTGCTG	del	1,704.0	interm	GR_E	160	1.00
				del		interm	GS_E	160	0.00
	7,370,701	AACCACCACCACCACCACACCACCACCACCACCACCACCT	AACCACCACCACCACCACACCACCACCACCACCACCACCT	ins	1,176.0	slow	GR_E	160	1.00
				ins		slow	GS_E	160	0.00
	7,605,600	TCACCACCACCTACCACCACACCACCACCAC	TCACCACCACCACCACCACACCACCACCAC	ins	1,133.0	interm	GR_D	160	1.00
				ins		interm	GS_D	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		CACCACCACCG	CACCACCACCACCG						
	7,708,394	GCCCCCCCCCCCCTCCCCCGGCCTGT	GCCCCCCCCCCCTCCCCCGGCCTGT	del	1,347.0	slow	GR_G	135	1.00
				del		slow	GS_G	135	0.00
17	147,544	TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTA	TTGTGTGTGTGTGTGTGTGTGTGTGTA	ins	2,019.0	slow	GR_E	160	0.96
				ins		slow	GS_E	160	0.00
	227,311	CACTTACTTA	CACTTA	del	1,313.0	interm	GR_D	160	1.00
				del		interm	GS_D	160	0.00
	235,339	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	AGTGTGTGTGTGTGTGTGTGTGTGTGTGTTA	ins	664.0	quick	GR_D	139	0.90
				ins		quick	GS_D	139	0.00
	841,961			ins	533.0	slow	GR_G	160	0.96

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		TCCACCACCACTCCACCACCAC CACCACCACC ACCACCACCA CCACCACCAC CACCACCACC ACCACCACCA CCACCACCAC CACG	TCCACCACCAC CACCACCACC ACCACCACCA CCACCACCAC CACCACCACC ACCACCACCA CCACCACCAC CACCACG	ins		slow	GS_G	160	0.00
	999,908	TTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGC	TTGTGTGTGTG TGTGTGTGTGT GTGTGTGTGTG TGTGTGTGTGT GTGTGC	ins	1,101.0	slow	GR_G	160	0.96
				ins		slow	GS_G	160	0.00
	1,957,433	GCCCCCCCCCC ATG	GCCCCCCCCCC CATG	ins	1,325.0	slow	GR_F	148	0.98
				ins		slow	GS_F	148	0.00
	2,488,437	GCACACACAC ACACACACAC	GCACACACAC ACACACACAC	del	1,103.0	quick	GR_E	141	1.00
				del		quick	GS_E	141	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		ACACACACAC ACACAT	ACACACACAC ACAT						
	2,823,518	C	A	snp	3,800.0	quick	GR_D	160	0.98
				snp		quick	GS_D	160	0.00
	3,843,005	CAGGGAGGGA GGGAGGGAGG GAGGGAGGGA GGGAGGGAGG CGAGGAAGGG AGGGAGG	CAGGGAGGGA GGGAGGGAGG GAGGGAGGGA GGGAGGGAGG GAGGCGAGGA AGGGAGGGAG G	ins	248.0	interm	GR_E	134	1.00
				ins		interm	GS_E	134	0.00
	4,891,587	G	T	snp	3,717.0	interm	GR_D	160	1.00
				snp		interm	GS_D	160	0.00
	4,968,828	GACACACACA CACACACACA CACACACACA CACATAC	GACACACACA CACACACACA CACACACACA CACACATAC	ins	971.0	slow	GR_G	137	1.00
				ins		slow	GS_G	137	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
	5,002,897	GTGGGGGGGGGGGTAAC	GTGGGGGGGGGGGGTAAC	ins	453.0	slow	GR_F	160	1.00
				ins		slow	GS_F	160	0.00
	5,452,301	TGGGGGGGCGGCA	TGGGGGGGGCGGCA	ins	758.0	interm	GR_E	141	1.00
				ins		interm	GS_E	141	0.00
	5,459,186	TTGTGTGTGTGTGTGTGTGTGTGGGTGTGTGTGTGTGTC	TTGTGTGTGTGTGTGTGTGTGGGTGTGTGTGTGTGTC	del	295.0	interm	GR_G	160	1.00
				del		interm	GS_G	160	0.00
	5,633,870	TTCCTCCTCCTCCTCCTCCTCCTCCTCTG	TTCCTCCTCCTCCTCCTCCTCCTCCTCTG	ins	1,063.0	interm	GR_D	160	1.00
				ins		interm	GS_D	160	0.00
18	531,627	CGTGTGTGGGTGTGTGGGTGTGGGTGT	GGTGTGTGGGTGTGTGGGTGTGGGTGT	snp	607.0	quick	GR_F	160	1.00
				snp		quick	GS_F	160	0.00

Chromosome	Position	Ref.allele	Alt.allele	Variant type	Variant calling quality	Selective history	Sample	GQ	AAF
		GGGTGTGTGGGTGTGTGGGTGTGA	GGGTGTGTGGGTGTGTGGGTGTGA						
	1,248,219	AAGCAGCAGCAGCAGCAGCA	AAGCAGCAGCAGCAGCAGCA	ins	1,906.0	slow	GR_F	160	1.00
		GCA	GCAGCA	ins		slow	GS_F	160	0.00

Appendix 6: Annotation of Variants susceptible to support glyphosate resistance from the VR experiment (Chapter5)

SnpEff (v4.3+T.galaxy2) Annotation of variants susceptible to support glyphosate resistance that are reported in Appendix 5: List of variants susceptible to support glyphosate resistance from the VR experiment (Chapter5)

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
01	292,515	Cre01.g001678	quick	4	3 prime UTR variant	MODIFIER
		Cre01.g001678	quick	4	3 prime UTR variant	MODIFIER
		Cre01.g001685	quick	4	upstream gene variant	MODIFIER
02	5,681,162	Cre02.g116750	quick	4	upstream gene variant	MODIFIER
		Cre02.g116850	quick	4	upstream gene variant	MODIFIER
		Cre02.g116750	quick	4	upstream gene variant	MODIFIER
		Cre02.g116800	quick	4	intron variant	MODIFIER
03	2,893,718	Cre03.g162950	quick	3	upstream gene variant	MODIFIER
		Cre03.g162900	quick	3	intron variant	MODIFIER
09	4,546,642	Cre09.g401293	quick	3	intron variant	MODIFIER
05	950,004	Cre05.g246900	interm	2	3 prime UTR variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre05.g246850	interm	2	upstream gene variant	MODIFIER
		Cre05.g246950	interm	2	upstream gene variant	MODIFIER
07	451,071	Cre07.g315350	slow	2	frameshift variant&splice region variant	HIGH
		Cre07.g315400	slow	2	upstream gene variant	MODIFIER
	668,524	Cre07.g317250	interm	2	3 prime UTR variant	MODIFIER
		Cre07.g317201	interm	2	upstream gene variant	MODIFIER
10	5,247,248	Cre10.g456400	interm	2	missense variant	MODERATE
		Cre10.g456400	interm	2	missense variant	MODERATE
		Cre10.g456350	interm	2	upstream gene variant	MODIFIER
15	1,850,364	Cre02.g141806	interm	2	upstream gene variant	MODIFIER
		Cre02.g141826	interm	2	intron variant	MODIFIER
01	217,144	Cre01.g001400	interm	1	upstream gene variant	MODIFIER
		Cre01.g001300	interm	1	downstream gene variant	MODIFIER
		Cre01.g001350	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	258,451	Cre01.g001650	slow	1	upstream gene variant	MODIFIER
		Cre01.g001657	slow	1	intron variant	MODIFIER
	430,725	Cre01.g002750	quick	1	upstream gene variant	MODIFIER
		Cre01.g002750	quick	1	upstream gene variant	MODIFIER
		Cre01.g002700	quick	1	intron variant	MODIFIER
	507,526	Cre01.g003050	interm	1	intron variant	MODIFIER
	587,370	Cre01.g003500	slow	1	disruptive inframe insertion	MODERATE
		Cre01.g003508	slow	1	upstream gene variant	MODIFIER
	754,962	Cre01.g004157	slow	1	upstream gene variant	MODIFIER
		Cre01.g004124	slow	1	downstream gene variant	MODIFIER
		Cre01.g004124	slow	1	downstream gene variant	MODIFIER
		Cre01.g004124-Cre01.g004157	slow	1	intergenic region	MODIFIER
	1,006,293	Cre01.g006150	quick	1	intron variant	MODIFIER
	1,282,990	Cre01.g006766	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	1,452,309	Cre01.g007774	slow	1	upstream gene variant	MODIFIER
		Cre01.g007811	slow	1	intron variant	MODIFIER
	1,488,267	Cre01.g008100	slow	1	3 prime UTR variant	MODIFIER
		Cre01.g008051	slow	1	downstream gene variant	MODIFIER
	1,551,285	Cre01.g008500	slow	1	upstream gene variant	MODIFIER
		Cre01.g008450	slow	1	intron variant	MODIFIER
	2,276,397	Cre01.g012850	slow	1	upstream gene variant	MODIFIER
		Cre01.g012750	slow	1	downstream gene variant	MODIFIER
		Cre01.g012800	slow	1	intron variant	MODIFIER
		Cre01.g012800	slow	1	intron variant	MODIFIER
	3,226,598	Cre01.g020800	quick	1	upstream gene variant	MODIFIER
		Cre01.g020850	quick	1	intron variant	MODIFIER
	3,631,914	Cre01.g023750	quick	1	upstream gene variant	MODIFIER
		Cre01.g023773	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre01.g023787	quick	1	intron variant	MODIFIER
	3,944,117	Cre01.g026016	interm	1	3 prime UTR variant	MODIFIER
		Cre01.g025983	interm	1	upstream gene variant	MODIFIER
		Cre01.g026050	interm	1	downstream gene variant	MODIFIER
	4,699,130	Cre01.g032900	interm	1	3 prime UTR variant	MODIFIER
		Cre01.g032900	interm	1	3 prime UTR variant	MODIFIER
		Cre01.g032900	interm	1	3 prime UTR variant	MODIFIER
		Cre01.g032950	interm	1	upstream gene variant	MODIFIER
		Cre01.g033000	interm	1	downstream gene variant	MODIFIER
	4,829,965	Cre01.g033900	interm	1	upstream gene variant	MODIFIER
		Cre01.g033900	interm	1	upstream gene variant	MODIFIER
		Cre01.g034000	interm	1	downstream gene variant	MODIFIER
		Cre01.g033950	interm	1	intron variant	MODIFIER
	5,510,782	Cre01.g039626	slow	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre01.g039600	slow	1	intron variant	MODIFIER
	5,636,548	Cre01.g040600	interm	1	missense variant	MODERATE
		Cre01.g040650	interm	1	downstream gene variant	MODIFIER
		Cre01.g040650	interm	1	downstream gene variant	MODIFIER
		Cre01.g040650	interm	1	downstream gene variant	MODIFIER
	6,196,146	Cre01.g044400	quick	1	upstream gene variant	MODIFIER
		Cre01.g044450	quick	1	downstream gene variant	MODIFIER
		Cre01.g044450-Cre01.g044550	quick	1	intergenic region	MODIFIER
	7,199,539	Cre01.g051625-Cre01.g051700	interm	1	intergenic region	MODIFIER
	7,226,873	Cre01.g051625-Cre01.g051700	interm	1	intergenic region	MODIFIER
	7,260,262	Cre01.g051625-Cre01.g051700	interm	1	intergenic region	MODIFIER
	7,386,055	Cre01.g052601	interm	1	intron variant	MODIFIER
	7,483,334	Cre01.g053150	interm	1	upstream gene variant	MODIFIER
		Cre01.g053150	interm	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre01.g053000	interm	1	downstream gene variant	MODIFIER
		Cre01.g053050	interm	1	downstream gene variant	MODIFIER
		Cre01.g053100	interm	1	intron variant	MODIFIER
	7,532,909	Cre01.g053350	interm	1	upstream gene variant	MODIFIER
		Cre01.g053360	interm	1	intron variant	MODIFIER
	8,062,539	Cre01.g067282	quick	1	upstream gene variant	MODIFIER
		Cre01.g068012	quick	1	upstream gene variant	MODIFIER
		Cre01.g067647	quick	1	downstream gene variant	MODIFIER
		Cre01.g067647-Cre01.g068012	quick	1	intergenic region	MODIFIER
02	344	CHR_START-Cre02.g073150	interm	1	intergenic region	MODIFIER
	360	CHR_START-Cre02.g073150	interm	1	intergenic region	MODIFIER
	978,214	Cre02.g079850	slow	1	downstream gene variant	MODIFIER
		Cre02.g079926	slow	1	intron variant	MODIFIER
	2,307,569	Cre02.g090900	interm	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre02.g090900	interm	1	downstream gene variant	MODIFIER
		Cre02.g090850	interm	1	intron variant	MODIFIER
	2,331,425	Cre02.g090950	slow	1	downstream gene variant	MODIFIER
		Cre02.g091000	slow	1	intron variant	MODIFIER
	2,658,571	Cre02.g093650	quick	1	downstream gene variant	MODIFIER
		Cre02.g093700	quick	1	downstream gene variant	MODIFIER
		Cre02.g093700	quick	1	downstream gene variant	MODIFIER
		Cre02.g093600	quick	1	intron variant	MODIFIER
	3,065,315	Cre02.g096300	quick	1	upstream gene variant	MODIFIER
		Cre02.g096400	quick	1	upstream gene variant	MODIFIER
		Cre02.g096300	quick	1	upstream gene variant	MODIFIER
		Cre02.g096350	quick	1	intron variant	MODIFIER
	3,073,896	Cre02.g096350	slow	1	downstream gene variant	MODIFIER
		Cre02.g096450	slow	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre02.g096455	slow	1	downstream gene variant	MODIFIER
		Cre02.g096400	slow	1	intron variant	MODIFIER
	3,438,348	Cre02.g099055	interm	1	downstream gene variant	MODIFIER
		Cre02.g099100	interm	1	downstream gene variant	MODIFIER
		Cre02.g099055-Cre02.g099100	interm	1	intergenic region	MODIFIER
	3,525,664	Cre02.g099601	quick	1	upstream gene variant	MODIFIER
		Cre02.g099700	quick	1	upstream gene variant	MODIFIER
		Cre02.g099650	quick	1	intron variant	MODIFIER
	3,762,754	Cre02.g101250	slow	1	upstream gene variant	MODIFIER
		Cre02.g101200	slow	1	intron variant	MODIFIER
	4,728,172	Cre02.g109100	quick	1	conservative inframe insertion	MODERATE
		Cre02.g109050	quick	1	downstream gene variant	MODIFIER
	6,054,586	Cre02.g119800	slow	1	synonymous variant	LOW
	6,162,955	Cre09.g387200	slow	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre09.g387150	slow	1	downstream gene variant	MODIFIER
		Cre09.g387150-Cre09.g387200	slow	1	intergenic region	MODIFIER
	6,214,981	Cre09.g387400	quick	1	intron variant	MODIFIER
	8,352,670	Cre02.g142400	quick	1	5 prime UTR variant	MODIFIER
		Cre02.g142605	quick	1	upstream gene variant	MODIFIER
		Cre02.g142350	quick	1	upstream gene variant	MODIFIER
		Cre02.g142606	quick	1	downstream gene variant	MODIFIER
	8,353,860	Cre02.g142350	interm	1	upstream gene variant	MODIFIER
		Cre02.g142606	interm	1	downstream gene variant	MODIFIER
		Cre02.g142400	interm	1	intron variant	MODIFIER
	8,409,942	Cre02.g142050	slow	1	upstream gene variant	MODIFIER
		Cre02.g141950	slow	1	upstream gene variant	MODIFIER
		Cre02.g142100	slow	1	downstream gene variant	MODIFIER
		Cre02.g142000	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
03	79,432	Cre03.g143787	quick	1	downstream gene variant	MODIFIER
		Cre03.g143787-Cre03.g143807	quick	1	intergenic region	MODIFIER
	79,751	Cre03.g143787	quick	1	downstream gene variant	MODIFIER
		Cre03.g143787-Cre03.g143807	quick	1	intergenic region	MODIFIER
	81,120	Cre03.g143787	quick	1	downstream gene variant	MODIFIER
		Cre03.g143787-Cre03.g143807	quick	1	intergenic region	MODIFIER
	226,029	Cre03.g144344	interm	1	downstream gene variant	MODIFIER
		Cre03.g144344	interm	1	downstream gene variant	MODIFIER
		Cre03.g144324	interm	1	intron variant	MODIFIER
	326,711	Cre03.g144627	quick	1	upstream gene variant	MODIFIER
		Cre03.g144607	quick	1	intron variant	MODIFIER
	470,045	Cre03.g145047	quick	1	upstream gene variant	MODIFIER
		Cre03.g145027	quick	1	intron variant	MODIFIER
	498,960	Cre03.g145127	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	514,072	Cre03.g145147	quick	1	upstream gene variant	MODIFIER
		Cre03.g145167	quick	1	downstream gene variant	MODIFIER
		Cre03.g145187	quick	1	intron variant	MODIFIER
	1,166,410	Cre03.g149450	slow	1	downstream gene variant	MODIFIER
		Cre03.g149400	slow	1	intron variant	MODIFIER
	1,445,531	Cre03.g151450	interm	1	upstream gene variant	MODIFIER
		Cre03.g151400	interm	1	intron variant	MODIFIER
	1,549,357	Cre03.g152425	slow	1	upstream gene variant	MODIFIER
		Cre03.g152450	slow	1	intron variant	MODIFIER
	1,894,897	Cre03.g155150	quick	1	conservative inframe insertion	MODERATE
		Cre03.g155100	quick	1	downstream gene variant	MODIFIER
		Cre03.g155200	quick	1	downstream gene variant	MODIFIER
	2,049,018	Cre03.g156600	slow	1	upstream gene variant	MODIFIER
		Cre03.g156500	slow	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre03.g156500	slow	1	downstream gene variant	MODIFIER
		Cre03.g156550	slow	1	intron variant	MODIFIER
	2,116,424	Cre03.g157050	interm	1	splice region variant&intron variant	LOW
		Cre03.g157100	interm	1	upstream gene variant	MODIFIER
		Cre03.g157150	interm	1	downstream gene variant	MODIFIER
	2,664,628	Cre03.g161150	slow	1	splice region variant&intron variant	LOW
		Cre03.g161100	slow	1	downstream gene variant	MODIFIER
	2,893,718	Cre03.g162950	interm	1	upstream gene variant	MODIFIER
		Cre03.g162900	interm	1	intron variant	MODIFIER
		Cre03.g162950	slow	1	upstream gene variant	MODIFIER
		Cre03.g162900	slow	1	intron variant	MODIFIER
	3,446,876	Cre03.g167351	slow	1	downstream gene variant	MODIFIER
		Cre03.g167450	slow	1	downstream gene variant	MODIFIER
		Cre03.g167400	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	4,374,908	Cre03.g175250	slow	1	synonymous variant	LOW
		Cre03.g175250	slow	1	synonymous variant	LOW
		Cre03.g175200	slow	1	upstream gene variant	MODIFIER
	4,957,186	Cre03.g179921	slow	1	upstream gene variant	MODIFIER
		Cre03.g179941	slow	1	downstream gene variant	MODIFIER
		Cre03.g179961	slow	1	intron variant	MODIFIER
	6,155,151	Cre03.g191900	interm	1	intron variant	MODIFIER
	7,571,962	Cre03.g203600	slow	1	3 prime UTR variant	MODIFIER
		Cre03.g203550	slow	1	upstream gene variant	MODIFIER
		Cre03.g203550	slow	1	upstream gene variant	MODIFIER
		Cre03.g203550	slow	1	upstream gene variant	MODIFIER
		Cre03.g203700	slow	1	downstream gene variant	MODIFIER
		Cre03.g203650	slow	1	downstream gene variant	MODIFIER
	8,769,394	Cre03.g204241	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	9,159,155	Cre03.g198975-Cre15.g634500	interm	1	intergenic region	MODIFIER
	9,185,690	Cre03.g198975-Cre15.g634500	slow	1	intergenic region	MODIFIER
04	54,403	Cre04.g214209	interm	1	downstream gene variant	MODIFIER
		Cre04.g214321	interm	1	intron variant	MODIFIER
	236,253	Cre04.g217200	quick	1	upstream gene variant	MODIFIER
		Cre04.g217400	quick	1	downstream gene variant	MODIFIER
		Cre04.g217200-Cre04.g217400	quick	1	intergenic region	MODIFIER
	1,182,986	Cre04.g215400	interm	1	intron variant	MODIFIER
	1,612,570	Cre04.g213000	slow	1	upstream gene variant	MODIFIER
		Cre04.g213002	slow	1	upstream gene variant	MODIFIER
		Cre04.g213000-Cre04.g213002	slow	1	intergenic region	MODIFIER
	1,624,635	Cre04.g212700	interm	1	intron variant	MODIFIER
	2,178,800	Cre04.g220150	interm	1	downstream gene variant	MODIFIER
		Cre04.g220200	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre04.g220200	interm	1	intron variant	MODIFIER
		Cre04.g220200	interm	1	intron variant	MODIFIER
	2,366,026	Cre04.g221550	interm	1	upstream gene variant	MODIFIER
		Cre04.g221450	interm	1	downstream gene variant	MODIFIER
		Cre04.g221500	interm	1	intron variant	MODIFIER
	2,514,063	Cre04.g222550	interm	1	missense variant	MODERATE
		Cre04.g222650	interm	1	upstream gene variant	MODIFIER
		Cre04.g222600	interm	1	downstream gene variant	MODIFIER
	2,988,151	Cre04.g225301	quick	1	upstream gene variant	MODIFIER
		Cre04.g225350	quick	1	downstream gene variant	MODIFIER
		Cre04.g225350	quick	1	downstream gene variant	MODIFIER
		Cre04.g225301-Cre04.g225350	quick	1	intergenic region	MODIFIER
	3,024,945	Cre04.g225700	slow	1	downstream gene variant	MODIFIER
		Cre04.g225750	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre04.g225750	slow	1	intron variant	MODIFIER
	3,251,425	Cre04.g227200	interm	1	downstream gene variant	MODIFIER
		Cre04.g227251	interm	1	intron variant	MODIFIER
	3,428,879	Cre04.g228100	interm	1	upstream gene variant	MODIFIER
		Cre04.g228050	interm	1	downstream gene variant	MODIFIER
		Cre04.g228000	interm	1	intron variant	MODIFIER
		Cre04.g228100	slow	1	upstream gene variant	MODIFIER
		Cre04.g228050	slow	1	downstream gene variant	MODIFIER
		Cre04.g228000	slow	1	intron variant	MODIFIER
05	230,918	Cre05.g243150	quick	1	missense variant	MODERATE
		Cre05.g243050	quick	1	upstream gene variant	MODIFIER
		Cre05.g243151	quick	1	upstream gene variant	MODIFIER
	429,545	Cre05.g243801	quick	1	intron variant	MODIFIER
		Cre05.g243801	quick	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre05.g243801	quick	1	intron variant	MODIFIER
	908,228	Cre05.g246550	slow	1	conservative inframe insertion	MODERATE
		Cre05.g246551	slow	1	downstream gene variant	MODIFIER
	1,004,177	Cre05.g247250	quick	1	upstream gene variant	MODIFIER
		Cre05.g247350	quick	1	upstream gene variant	MODIFIER
		Cre05.g247300	quick	1	intron variant	MODIFIER
	1,137,275	Cre05.g248401	quick	1	conservative inframe insertion	MODERATE
		Cre05.g248300	quick	1	downstream gene variant	MODIFIER
		Cre05.g248400	quick	1	downstream gene variant	MODIFIER
		Cre05.g248401	quick	1	downstream gene variant	MODIFIER
	2,130,722	Cre05.g237000	slow	1	3 prime UTR variant	MODIFIER
		Cre05.g236950	slow	1	upstream gene variant	MODIFIER
		Cre05.g237050	slow	1	upstream gene variant	MODIFIER
	2,297,133	Cre05.g238250	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre05.g238260	quick	1	intron variant	MODIFIER
	2,357,694	Cre05.g238374	quick	1	intron variant	MODIFIER
	2,620,554	Cre05.g240225	interm	1	intron variant	MODIFIER
		Cre05.g240225	interm	1	intron variant	MODIFIER
	2,735,633	Cre05.g240850	quick	1	downstream gene variant	MODIFIER
		Cre05.g240900	quick	1	intron variant	MODIFIER
		Cre05.g240900	quick	1	intron variant	MODIFIER
	2,756,792	Cre05.g241100	interm	1	upstream gene variant	MODIFIER
		Cre05.g241050	interm	1	downstream gene variant	MODIFIER
		Cre05.g241150	interm	1	intron variant	MODIFIER
	2,990,559	Cre05.g232004	slow	1	downstream gene variant	MODIFIER
		Cre05.g232004-Cre24.g755397	slow	1	intergenic region	MODIFIER
06	307,123	Cre06.g250902	quick	1	upstream gene variant	MODIFIER
		Cre06.g250950	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre06.g250976	quick	1	downstream gene variant	MODIFIER
		Cre06.g250902-Cre06.g250950	quick	1	intergenic region	MODIFIER
	1,971,272	Cre06.g263750	interm	1	downstream gene variant	MODIFIER
		Cre06.g263800	interm	1	intron variant	MODIFIER
	2,462,667	Cre06.g268150	quick	1	upstream gene variant	MODIFIER
		Cre06.g268228	quick	1	downstream gene variant	MODIFIER
		Cre06.g268200	quick	1	intron variant	MODIFIER
	3,530,061	Cre06.g278103	slow	1	3 prime UTR variant	MODIFIER
		Cre06.g278105	slow	1	upstream gene variant	MODIFIER
		Cre06.g278104	slow	1	downstream gene variant	MODIFIER
	4,075,824	Cre06.g278208	interm	1	downstream gene variant	MODIFIER
		Cre06.g278209	interm	1	downstream gene variant	MODIFIER
		Cre06.g278209	interm	1	downstream gene variant	MODIFIER
		Cre06.g278208-Cre06.g278209	interm	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	4,599,211	Cre06.g278296	quick	1	5 prime UTR variant	MODIFIER
		Cre06.g278278	quick	1	upstream gene variant	MODIFIER
		Cre06.g278297	quick	1	upstream gene variant	MODIFIER
		Cre06.g278295	quick	1	downstream gene variant	MODIFIER
	5,097,107	Cre06.g282100	slow	1	upstream gene variant	MODIFIER
		Cre06.g282200	slow	1	upstream gene variant	MODIFIER
		Cre06.g282250	slow	1	upstream gene variant	MODIFIER
		Cre06.g282050	slow	1	downstream gene variant	MODIFIER
		Cre06.g282150	slow	1	intron variant	MODIFIER
	5,480,613	Cre06.g285650	slow	1	upstream gene variant	MODIFIER
		Cre06.g285700	slow	1	intron variant	MODIFIER
	5,711,460	Cre06.g287800	quick	1	upstream gene variant	MODIFIER
		Cre06.g287700	quick	1	downstream gene variant	MODIFIER
		Cre06.g287750	quick	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	6,375,099	Cre06.g292950	quick	1	upstream gene variant	MODIFIER
		Cre06.g292850	quick	1	downstream gene variant	MODIFIER
		Cre06.g292900	quick	1	intron variant	MODIFIER
	7,318,185	Cre06.g299500	interm	1	upstream gene variant	MODIFIER
		Cre06.g299650	interm	1	upstream gene variant	MODIFIER
		Cre06.g299550	interm	1	downstream gene variant	MODIFIER
		Cre06.g299600	interm	1	intron variant	MODIFIER
	7,529,017	Cre17.g719813	quick	1	upstream gene variant	MODIFIER
		Cre06.g301450	quick	1	upstream gene variant	MODIFIER
		Cre06.g301251	quick	1	downstream gene variant	MODIFIER
		Cre06.g301251-Cre17.g719813	quick	1	intergenic region	MODIFIER
	7,633,103	Cre06.g302050	quick	1	intron variant	MODIFIER
	8,237,852	Cre06.g306700	interm	1	upstream gene variant	MODIFIER
		Cre06.g306601	interm	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre06.g306650	interm	1	intron variant	MODIFIER
		Cre06.g306650	interm	1	intron variant	MODIFIER
	8,629,983	Cre06.g309951	slow	1	intron variant	MODIFIER
07	21,924	Cre07.g312002	quick	1	downstream gene variant	MODIFIER
		Cre07.g312002-Cre07.g312050	quick	1	intergenic region	MODIFIER
	24,590	Cre07.g312050	quick	1	downstream gene variant	MODIFIER
		Cre07.g312002-Cre07.g312050	quick	1	intergenic region	MODIFIER
	25,017	Cre07.g312050	quick	1	downstream gene variant	MODIFIER
		Cre07.g312002-Cre07.g312050	quick	1	intergenic region	MODIFIER
	189,532	Cre07.g313350	interm	1	downstream gene variant	MODIFIER
		Cre07.g313450	interm	1	downstream gene variant	MODIFIER
		Cre07.g313400	interm	1	intron variant	MODIFIER
		Cre07.g313400	interm	1	intron variant	MODIFIER
	359,246	Cre07.g314700	slow	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre07.g314600	slow	1	downstream gene variant	MODIFIER
		Cre07.g314676	slow	1	downstream gene variant	MODIFIER
		Cre07.g314650	slow	1	intron variant	MODIFIER
	640,757	Cre07.g316992	slow	1	upstream gene variant	MODIFIER
		Cre07.g316992	slow	1	intron variant	MODIFIER
		Cre07.g316992	slow	1	intron variant	MODIFIER
		Cre07.g316992	slow	1	intron variant	MODIFIER
		Cre07.g316992	slow	1	intron variant	MODIFIER
	1,399,520	Cre07.g323000	interm	1	downstream gene variant	MODIFIER
		Cre07.g323050	interm	1	downstream gene variant	MODIFIER
		Cre07.g323100	interm	1	intron variant	MODIFIER
	1,994,476	Cre07.g325761	quick	1	intron variant	MODIFIER
	2,130,886	Cre07.g326626	interm	1	conservative inframe insertion	MODERATE
		Cre07.g326650	interm	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre07.g326700	interm	1	upstream gene variant	MODIFIER
		Cre07.g326600	interm	1	downstream gene variant	MODIFIER
	2,234,373	Cre07.g327600	slow	1	splice region variant&intron variant	LOW
		Cre07.g327600	slow	1	splice region variant&intron variant	LOW
		Cre07.g327550	slow	1	upstream gene variant	MODIFIER
		Cre07.g327650	slow	1	downstream gene variant	MODIFIER
	2,307,693	Cre07.g328000	slow	1	upstream gene variant	MODIFIER
		Cre07.g328075	slow	1	upstream gene variant	MODIFIER
		Cre07.g328075	slow	1	upstream gene variant	MODIFIER
		Cre07.g328075	slow	1	upstream gene variant	MODIFIER
		Cre07.g328050	slow	1	intron variant	MODIFIER
	2,412,917	Cre07.g328950	quick	1	upstream gene variant	MODIFIER
		Cre07.g329050	quick	1	downstream gene variant	MODIFIER
		Cre07.g329000	quick	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	2,674,272	Cre07.g330900	interm	1	upstream gene variant	MODIFIER
		Cre07.g330950	interm	1	upstream gene variant	MODIFIER
		Cre07.g331050	interm	1	upstream gene variant	MODIFIER
		Cre07.g331050	interm	1	upstream gene variant	MODIFIER
		Cre07.g331000	interm	1	intron variant	MODIFIER
	2,960,519	Cre07.g333150	quick	1	upstream gene variant	MODIFIER
		Cre07.g333100	quick	1	downstream gene variant	MODIFIER
		Cre07.g333350	quick	1	downstream gene variant	MODIFIER
		Cre07.g333100	quick	1	downstream gene variant	MODIFIER
		Cre07.g333150-Cre07.g333350	quick	1	intergenic region	MODIFIER
	3,324,261	Cre07.g335050	interm	1	upstream gene variant	MODIFIER
		Cre07.g335000	interm	1	downstream gene variant	MODIFIER
		Cre07.g335000	interm	1	downstream gene variant	MODIFIER
		Cre07.g335000-Cre07.g335050	interm	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	3,955,178	Cre07.g339926	quick	1	upstream gene variant	MODIFIER
		Cre07.g339950	quick	1	upstream gene variant	MODIFIER
		Cre07.g339950	quick	1	upstream gene variant	MODIFIER
		Cre07.g339900	quick	1	downstream gene variant	MODIFIER
		Cre07.g340000	quick	1	downstream gene variant	MODIFIER
		Cre07.g339926-Cre07.g339950	quick	1	intergenic region	MODIFIER
	4,174,874	Cre07.g341925	slow	1	upstream gene variant	MODIFIER
		Cre07.g342000	slow	1	downstream gene variant	MODIFIER
		Cre07.g341950	slow	1	intron variant	MODIFIER
	4,783,188	Cre07.g346000	slow	1	intron variant	MODIFIER
	4,848,576	Cre07.g346400	interm	1	upstream gene variant	MODIFIER
		Cre07.g346450	interm	1	upstream gene variant	MODIFIER
		Cre07.g346500	interm	1	upstream gene variant	MODIFIER
		Cre07.g346418	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	5,054,903	Cre07.g347980	slow	1	upstream gene variant	MODIFIER
		Cre07.g347980	slow	1	intron variant	MODIFIER
		Cre07.g348010	slow	1	intron variant	MODIFIER
	5,496,794	Cre07.g351550	slow	1	upstream gene variant	MODIFIER
		Cre07.g351650	slow	1	upstream gene variant	MODIFIER
		Cre07.g351600	slow	1	intron variant	MODIFIER
	5,975,301	Cre07.g354900	interm	1	upstream gene variant	MODIFIER
		Cre07.g354850	interm	1	intron variant	MODIFIER
	6,095,652	Cre07.g356050	slow	1	upstream gene variant	MODIFIER
		Cre07.g355950	slow	1	downstream gene variant	MODIFIER
		Cre07.g356000	slow	1	intron variant	MODIFIER
	6,251,047	Cre07.g356970	quick	1	upstream gene variant	MODIFIER
		Cre07.g356970	quick	1	upstream gene variant	MODIFIER
		Cre07.g356980	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre07.g356960	quick	1	intron variant	MODIFIER
	6,254,608	Cre07.g356960	quick	1	upstream gene variant	MODIFIER
		Cre07.g357000	quick	1	upstream gene variant	MODIFIER
		Cre07.g356970	quick	1	downstream gene variant	MODIFIER
		Cre07.g356970	quick	1	downstream gene variant	MODIFIER
		Cre07.g356980	quick	1	intron variant	MODIFIER
08	416,680	Cre08.g358650	interm	1	upstream gene variant	MODIFIER
		Cre08.g358616	interm	1	downstream gene variant	MODIFIER
		Cre08.g358600	interm	1	intron variant	MODIFIER
	496,202	Cre08.g358900	slow	1	upstream gene variant	MODIFIER
		Cre08.g359000	slow	1	downstream gene variant	MODIFIER
		Cre08.g359000	slow	1	downstream gene variant	MODIFIER
		Cre08.g358950	slow	1	intron variant	MODIFIER
		Cre08.g358950	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	581,982	Cre08.g359650	quick	1	3 prime UTR variant	MODIFIER
		Cre08.g359600	quick	1	downstream gene variant	MODIFIER
		Cre08.g359650	interm	1	3 prime UTR variant	MODIFIER
		Cre08.g359600	interm	1	downstream gene variant	MODIFIER
	594,711	Cre08.g359700	interm	1	upstream gene variant	MODIFIER
		Cre08.g359800	interm	1	downstream gene variant	MODIFIER
		Cre08.g359750	interm	1	intron variant	MODIFIER
		Cre08.g359750	interm	1	intron variant	MODIFIER
	720,914	Cre08.g360500	quick	1	upstream gene variant	MODIFIER
		Cre08.g360450	quick	1	intron variant	MODIFIER
	1,334,262	Cre08.g363950	slow	1	upstream gene variant	MODIFIER
		Cre08.g363874	slow	1	downstream gene variant	MODIFIER
		Cre08.g363874	slow	1	downstream gene variant	MODIFIER
		Cre08.g363874-Cre08.g363950	slow	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	1,958,906	Cre08.g368100	slow	1	conservative inframe insertion	MODERATE
		Cre08.g368050	slow	1	upstream gene variant	MODIFIER
		Cre08.g368150	slow	1	downstream gene variant	MODIFIER
	2,940,149	Cre08.g374600	interm	1	upstream gene variant	MODIFIER
		Cre08.g374700	interm	1	upstream gene variant	MODIFIER
		Cre08.g374650	interm	1	intron variant	MODIFIER
	3,214,279	Cre08.g376740	slow	1	upstream gene variant	MODIFIER
		Cre08.g376740	slow	1	upstream gene variant	MODIFIER
		Cre33.g758897	slow	1	downstream gene variant	MODIFIER
		Cre33.g758897	slow	1	downstream gene variant	MODIFIER
		Cre08.g376720	slow	1	intron variant	MODIFIER
		Cre08.g376720	slow	1	intron variant	MODIFIER
	3,421,934	Cre08.g378050	quick	1	downstream gene variant	MODIFIER
		Cre08.g378100	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre08.g378150	quick	1	intron variant	MODIFIER
	3,460,230	Cre08.g378450	interm	1	3 prime UTR variant	MODIFIER
		Cre08.g378417	interm	1	upstream gene variant	MODIFIER
		Cre08.g378417	interm	1	upstream gene variant	MODIFIER
		Cre08.g378500	interm	1	downstream gene variant	MODIFIER
	3,557,195	Cre08.g379300	interm	1	intron variant	MODIFIER
	3,948,362	Cre08.g382250	quick	1	upstream gene variant	MODIFIER
		Cre08.g382200	quick	1	intron variant	MODIFIER
		Cre08.g382200	quick	1	intron variant	MODIFIER
	4,031,353	Cre08.g382620	interm	1	intron variant	MODIFIER
	4,197,249	Cre08.g383750	slow	1	3 prime UTR variant	MODIFIER
		Cre08.g383750	slow	1	3 prime UTR variant	MODIFIER
		Cre08.g383800	slow	1	upstream gene variant	MODIFIER
		Cre08.g383702	slow	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	4,397,893	Cre08.g385000	interm	1	3 prime UTR variant	MODIFIER
		Cre08.g385050	interm	1	upstream gene variant	MODIFIER
09	1,209,544	Cre09.g400950	interm	1	frameshift variant	HIGH
		Cre09.g401000	interm	1	downstream gene variant	MODIFIER
		Cre09.g400900	interm	1	downstream gene variant	MODIFIER
	1,216,520	Cre09.g400900	interm	1	3 prime UTR variant	MODIFIER
		Cre09.g400950	interm	1	downstream gene variant	MODIFIER
	1,417,560	Cre09.g399950	slow	1	downstream gene variant	MODIFIER
		Cre09.g400000	slow	1	intron variant	MODIFIER
	1,579,392	Cre09.g399073	slow	1	downstream gene variant	MODIFIER
		Cre09.g399000	slow	1	downstream gene variant	MODIFIER
		Cre09.g399050	slow	1	intron variant	MODIFIER
	1,710,470	Cre09.g398556	interm	1	upstream gene variant	MODIFIER
		Cre09.g398554	interm	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre09.g398555	interm	1	downstream gene variant	MODIFIER
		Cre09.g398500	interm	1	downstream gene variant	MODIFIER
		Cre09.g398555-Cre09.g398556	interm	1	intergenic region	MODIFIER
	3,816,199	Cre09.g397845	quick	1	conservative inframe insertion	MODERATE
		Cre09.g397808	quick	1	upstream gene variant	MODIFIER
	3,856,765	Cre09.g397919	quick	1	downstream gene variant	MODIFIER
		Cre09.g397956	quick	1	intron variant	MODIFIER
	4,002,402	Cre09.g398771	quick	1	upstream gene variant	MODIFIER
		Cre09.g398808	quick	1	downstream gene variant	MODIFIER
		Cre09.g398734-Cre09.g398771	quick	1	intergenic region	MODIFIER
	4,648,989	Cre09.g401886	interm	1	upstream gene variant	MODIFIER
		Cre09.g401923	interm	1	intron variant	MODIFIER
	4,648,994	Cre09.g401886	interm	1	upstream gene variant	MODIFIER
		Cre09.g401923	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	4,688,496	Cre09.g402108	quick	1	intron variant	MODIFIER
	4,854,741	Cre09.g403071	slow	1	3 prime UTR variant	MODIFIER
		Cre09.g402997	slow	1	upstream gene variant	MODIFIER
		Cre09.g403108	slow	1	upstream gene variant	MODIFIER
		Cre09.g403034	slow	1	downstream gene variant	MODIFIER
	5,411,399	Cre09.g408550	slow	1	upstream gene variant	MODIFIER
		Cre09.g408464	slow	1	downstream gene variant	MODIFIER
		Cre09.g408500	slow	1	intron variant	MODIFIER
	5,413,779	Cre09.g408550	quick	1	upstream gene variant	MODIFIER
		Cre09.g408464	quick	1	downstream gene variant	MODIFIER
		Cre09.g408500	quick	1	intron variant	MODIFIER
	5,418,408	Cre09.g408500	quick	1	upstream gene variant	MODIFIER
		Cre09.g408600	quick	1	upstream gene variant	MODIFIER
		Cre09.g408550	quick	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	5,524,304	Cre09.g409325	slow	1	3 prime UTR variant	MODIFIER
		Cre09.g409300	slow	1	downstream gene variant	MODIFIER
		Cre09.g409350	slow	1	downstream gene variant	MODIFIER
	5,569,262	Cre09.g409700	slow	1	upstream gene variant	MODIFIER
		Cre09.g409700-Cre09.g409728	slow	1	intergenic region	MODIFIER
	5,712,694	Cre09.g410500	interm	1	upstream gene variant	MODIFIER
		Cre09.g410500	interm	1	upstream gene variant	MODIFIER
		Cre09.g410650	interm	1	upstream gene variant	MODIFIER
		Cre09.g410600	interm	1	downstream gene variant	MODIFIER
		Cre09.g410650	interm	1	intron variant	MODIFIER
		Cre09.g410650	interm	1	intron variant	MODIFIER
	5,926,830	Cre09.g412150	interm	1	downstream gene variant	MODIFIER
		Cre09.g412175	interm	1	intron variant	MODIFIER
	6,143,181	Cre09.g413141-Cre09.g413150	interm	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	6,210,316	Cre09.g413533	interm	1	3 prime UTR variant	MODIFIER
		Cre09.g413566	interm	1	upstream gene variant	MODIFIER
		Cre09.g413500	interm	1	downstream gene variant	MODIFIER
	6,426,674	Cre09.g415300	quick	1	upstream gene variant	MODIFIER
		Cre09.g415350	quick	1	upstream gene variant	MODIFIER
		Cre09.g415400	quick	1	upstream gene variant	MODIFIER
		Cre09.g415250	quick	1	downstream gene variant	MODIFIER
		Cre09.g415300-Cre09.g415350	quick	1	intergenic region	MODIFIER
10	388,358	Cre10.g420450	quick	1	intron variant	MODIFIER
	660,244	Cre10.g421900	quick	1	upstream gene variant	MODIFIER
		Cre10.g421950	quick	1	downstream gene variant	MODIFIER
		Cre10.g421950	quick	1	downstream gene variant	MODIFIER
		Cre10.g421900-Cre10.g421950	quick	1	intergenic region	MODIFIER
	965,483	Cre10.g424450	interm	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre10.g424500	interm	1	downstream gene variant	MODIFIER
		Cre10.g424550	interm	1	intron variant	MODIFIER
	1,564,912	Cre10.g429000	quick	1	downstream gene variant	MODIFIER
		Cre10.g429000	quick	1	downstream gene variant	MODIFIER
		Cre10.g428966	quick	1	intron variant	MODIFIER
	1,777,785	Cre10.g430650	quick	1	upstream gene variant	MODIFIER
		Cre10.g430700	quick	1	intron variant	MODIFIER
	1,938,590	Cre10.g431900	interm	1	5 prime UTR variant	MODIFIER
		Cre10.g431850	interm	1	upstream gene variant	MODIFIER
		Cre10.g431950	interm	1	upstream gene variant	MODIFIER
		Cre10.g431800	interm	1	downstream gene variant	MODIFIER
	2,113,406	Cre10.g433100	slow	1	upstream gene variant	MODIFIER
		Cre10.g433200	slow	1	downstream gene variant	MODIFIER
		Cre10.g433150	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	2,534,998	Cre10.g436200	quick	1	frameshift variant	HIGH
	2,932,776	Cre10.g440000	interm	1	splice donor variant&intron variant	HIGH
		Cre10.g439950	interm	1	upstream gene variant	MODIFIER
	3,226,288	Cre10.g442300	slow	1	upstream gene variant	MODIFIER
		Cre10.g442350	slow	1	upstream gene variant	MODIFIER
		Cre10.g442450	slow	1	upstream gene variant	MODIFIER
		Cre10.g442300	slow	1	upstream gene variant	MODIFIER
		Cre10.g442400	slow	1	downstream gene variant	MODIFIER
		Cre10.g442500	slow	1	downstream gene variant	MODIFIER
		Cre10.g442400-Cre10.g442450	slow	1	intergenic region	MODIFIER
	3,285,917	Cre10.g442950	slow	1	downstream gene variant	MODIFIER
		Cre10.g443000	slow	1	intron variant	MODIFIER
	4,329,337	Cre10.g449750	slow	1	upstream gene variant	MODIFIER
		Cre10.g449850	slow	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre10.g449800	slow	1	intron variant	MODIFIER
	4,425,835	Cre10.g450626	slow	1	upstream gene variant	MODIFIER
		Cre10.g450550	slow	1	downstream gene variant	MODIFIER
		Cre10.g450600	slow	1	intron variant	MODIFIER
	4,731,857	Cre10.g452550	interm	1	upstream gene variant	MODIFIER
		Cre10.g452500	interm	1	intron variant	MODIFIER
	4,840,394	Cre10.g453450	interm	1	upstream gene variant	MODIFIER
		Cre10.g453400	interm	1	intron variant	MODIFIER
	5,047,265	Cre10.g454951	interm	1	downstream gene variant	MODIFIER
		Cre10.g454900	interm	1	intron variant	MODIFIER
	5,246,752	Cre10.g456400	interm	1	missense variant&splice region variant	MODERATE
		Cre10.g456400	interm	1	missense variant&splice region variant	MODERATE
		Cre10.g456350	interm	1	upstream gene variant	MODIFIER
	5,647,470	Cre10.g459300	interm	1	conservative inframe insertion	MODERATE

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre10.g459226	interm	1	upstream gene variant	MODIFIER
		Cre10.g459250	interm	1	upstream gene variant	MODIFIER
		Cre10.g459350	interm	1	downstream gene variant	MODIFIER
	5,660,729	Cre10.g459350	interm	1	upstream gene variant	MODIFIER
		Cre10.g459500	interm	1	upstream gene variant	MODIFIER
		Cre10.g459400	interm	1	downstream gene variant	MODIFIER
		Cre10.g459450	interm	1	intron variant	MODIFIER
	5,831,355	Cre10.g460750	interm	1	upstream gene variant	MODIFIER
		Cre10.g460650	interm	1	downstream gene variant	MODIFIER
		Cre10.g460700	interm	1	downstream gene variant	MODIFIER
		Cre10.g460700-Cre10.g460750	interm	1	intergenic region	MODIFIER
	6,137,866	Cre10.g462850	slow	1	conservative inframe deletion	MODERATE
		Cre10.g462816	slow	1	downstream gene variant	MODIFIER
	6,401,976	Cre10.g464950	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	6,475,403	Cre10.g465400	quick	1	downstream gene variant	MODIFIER
		Cre10.g465450	quick	1	intron variant	MODIFIER
	6,719,206	Cre10.g467200-CHR_END	interm	1	intergenic region	MODIFIER
11	950	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	965	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	972	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	984	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	1,008	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	1,033	CHR_START-Cre11.g467522	slow	1	intergenic region	MODIFIER
	1,296,346	Cre11.g467706	slow	1	upstream gene variant	MODIFIER
		Cre11.g467707	slow	1	upstream gene variant	MODIFIER
		Cre11.g467709	slow	1	downstream gene variant	MODIFIER
		Cre11.g467708	slow	1	intron variant	MODIFIER
		Cre11.g467708	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	1,473,139	Cre11.g468700	interm	1	upstream gene variant	MODIFIER
		Cre11.g468850	interm	1	downstream gene variant	MODIFIER
		Cre11.g468800	interm	1	downstream gene variant	MODIFIER
		Cre11.g468750	interm	1	intron variant	MODIFIER
	1,813,533	Cre11.g467750	quick	1	stop gained	HIGH
		Cre11.g467700	quick	1	upstream gene variant	MODIFIER
		Cre11.g467800	quick	1	downstream gene variant	MODIFIER
	2,148,461	Cre11.g467746	interm	1	downstream gene variant	MODIFIER
		Cre11.g467745	interm	1	intron variant	MODIFIER
	2,436,547	Cre02.g095125	slow	1	intron variant	MODIFIER
	3,290,360	Cre11.g476250	slow	1	upstream gene variant	MODIFIER
		Cre11.g476300	slow	1	intron variant	MODIFIER
	3,304,362	Cre49.g761297	quick	1	downstream gene variant	MODIFIER
		Cre11.g476300-Cre49.g761297	quick	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	3,365,906	Cre11.g476650	interm	1	intron variant	MODIFIER
	4,229,438	Cre11.g481500	quick	1	upstream gene variant	MODIFIER
		Cre11.g481550	quick	1	intron variant	MODIFIER
	4,299,599	Cre11.g482050	interm	1	upstream gene variant	MODIFIER
		Cre11.g482101	interm	1	intron variant	MODIFIER
12	137,918	Cre12.g484250	interm	1	downstream gene variant	MODIFIER
		Cre12.g484300	interm	1	downstream gene variant	MODIFIER
		Cre12.g484350	interm	1	intron variant	MODIFIER
	702,175	Cre12.g492750	slow	1	splice acceptor variant&intron variant	HIGH
		Cre12.g492851	slow	1	upstream gene variant	MODIFIER
	829,202	Cre12.g493404	slow	1	intron variant	MODIFIER
	829,232	Cre12.g493404	slow	1	intron variant	MODIFIER
	1,111,229	Cre12.g490891	interm	1	downstream gene variant	MODIFIER
		Cre12.g490850	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	1,119,261	Cre12.g490800	slow	1	downstream gene variant	MODIFIER
		Cre12.g490850	slow	1	intron variant	MODIFIER
	1,168,724	Cre12.g490350	quick	1	upstream gene variant	MODIFIER
		Cre12.g490250	quick	1	downstream gene variant	MODIFIER
		Cre12.g490300	quick	1	intron variant	MODIFIER
	1,653,521	Cre12.g486350	slow	1	intron variant	MODIFIER
	1,653,565	Cre12.g486350	slow	1	intron variant	MODIFIER
	2,748,952	Cre12.g503700	quick	1	upstream gene variant	MODIFIER
		Cre12.g503750	quick	1	intron variant	MODIFIER
	2,999,164	Cre12.g501450	interm	1	downstream gene variant	MODIFIER
		Cre12.g501403	interm	1	intron variant	MODIFIER
	3,272,417	Cre12.g498700	quick	1	upstream gene variant	MODIFIER
		Cre12.g498650	quick	1	upstream gene variant	MODIFIER
		Cre12.g498750	quick	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	3,768,419	Cre12.g514850	interm	1	upstream gene variant	MODIFIER
		Cre12.g514750	interm	1	downstream gene variant	MODIFIER
		Cre12.g514800	interm	1	intron variant	MODIFIER
	4,370,490	Cre12.g520050	slow	1	upstream gene variant	MODIFIER
		Cre12.g520000	slow	1	downstream gene variant	MODIFIER
		Cre12.g520072	slow	1	downstream gene variant	MODIFIER
		Cre12.g520100	slow	1	downstream gene variant	MODIFIER
		Cre12.g520000-Cre12.g520050	slow	1	intergenic region	MODIFIER
	4,618,179	Cre12.g522400	interm	1	3 prime UTR variant	MODIFIER
		Cre12.g522450	interm	1	upstream gene variant	MODIFIER
	6,222,214	Cre12.g536200	interm	1	5 prime UTR variant	MODIFIER
		Cre12.g536251	interm	1	downstream gene variant	MODIFIER
	6,328,210	Cre12.g537000	quick	1	downstream gene variant	MODIFIER
		Cre12.g537050	quick	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre12.g537100	quick	1	intron variant	MODIFIER
	6,336,302	Cre12.g537100	interm	1	intron variant	MODIFIER
		Cre12.g537100	slow	1	intron variant	MODIFIER
	6,575,282	Cre12.g538650	quick	1	3 prime UTR variant	MODIFIER
		Cre12.g538600	quick	1	upstream gene variant	MODIFIER
		Cre12.g538700	quick	1	upstream gene variant	MODIFIER
	9,088,322	Cre12.g542950	interm	1	upstream gene variant	MODIFIER
		Cre12.g542900	interm	1	intron variant	MODIFIER
13	233,788	Cre13.g563600	interm	1	downstream gene variant	MODIFIER
		Cre13.g563550	interm	1	intron variant	MODIFIER
	264,383	Cre13.g563733	slow	1	3 prime UTR variant	MODIFIER
		Cre13.g563733	slow	1	3 prime UTR variant	MODIFIER
		Cre13.g563800	slow	1	upstream gene variant	MODIFIER
		Cre13.g563700	slow	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	518,989	Cre13.g565260	interm	1	intron variant	MODIFIER
	811,394	Cre13.g567200	interm	1	intron variant	MODIFIER
		Cre13.g567200	interm	1	intron variant	MODIFIER
	2,354,166	Cre13.g579150	interm	1	upstream gene variant	MODIFIER
		Cre13.g579200	interm	1	upstream gene variant	MODIFIER
		Cre13.g579100	interm	1	intron variant	MODIFIER
	2,576,026	Cre13.g580650	interm	1	upstream gene variant	MODIFIER
		Cre13.g580750	interm	1	upstream gene variant	MODIFIER
		Cre13.g580700	interm	1	intron variant	MODIFIER
	3,088,622	Cre13.g584600	interm	1	upstream gene variant	MODIFIER
		Cre13.g584650	interm	1	downstream gene variant	MODIFIER
		Cre13.g584619	interm	1	intron variant	MODIFIER
	3,343,294	Cre13.g586600	interm	1	missense variant	MODERATE
		Cre13.g586500	interm	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre13.g586550	interm	1	upstream gene variant	MODIFIER
		Cre13.g586650	interm	1	upstream gene variant	MODIFIER
	3,768,290	Cre13.g589100	slow	1	upstream gene variant	MODIFIER
		Cre13.g589050	slow	1	intron variant	MODIFIER
	3,851,184	Cre13.g589600	interm	1	3 prime UTR variant	MODIFIER
		Cre17.g701884	interm	1	upstream gene variant	MODIFIER
		Cre13.g597676	interm	1	downstream gene variant	MODIFIER
	4,094,783	Cre13.g591073	slow	1	conservative inframe deletion	MODERATE
		Cre13.g591100	slow	1	downstream gene variant	MODIFIER
	4,638,826	Cre13.g603950	slow	1	missense variant	MODERATE
		Cre13.g603900	slow	1	upstream gene variant	MODIFIER
		Cre13.g604000	slow	1	upstream gene variant	MODIFIER
	4,693,224	Cre13.g604250	interm	1	downstream gene variant	MODIFIER
		Cre13.g604200	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	4,794,827	Cre13.g604905	slow	1	3 prime UTR variant	MODIFIER
		Cre13.g604850	slow	1	downstream gene variant	MODIFIER
		Cre13.g604950	slow	1	downstream gene variant	MODIFIER
		Cre13.g604950	slow	1	downstream gene variant	MODIFIER
		Cre13.g604950	slow	1	downstream gene variant	MODIFIER
	5,163,079	Cre13.g607450	interm	1	upstream gene variant	MODIFIER
		Cre13.g607400	interm	1	intron variant	MODIFIER
	5,177,388	Cre13.g607550	slow	1	downstream gene variant	MODIFIER
		Cre13.g607500	slow	1	intron variant	MODIFIER
	5,177,487	Cre13.g607550	interm	1	downstream gene variant	MODIFIER
		Cre13.g607500	interm	1	intron variant	MODIFIER
14	208	CHR_START-Cre14.g608050	slow	1	intergenic region	MODIFIER
	311	CHR_START-Cre14.g608050	slow	1	intergenic region	MODIFIER
	330	CHR_START-Cre14.g608050	slow	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	227,235	Cre14.g609202	quick	1	intron variant	MODIFIER
	1,448,634	Cre14.g617700	slow	1	intron variant	MODIFIER
	1,603,279	Cre14.g618750	interm	1	splice region variant&intron variant	LOW
		Cre14.g618776	interm	1	upstream gene variant	MODIFIER
	1,651,829	Cre14.g618926	quick	1	disruptive inframe deletion	MODERATE
		Cre14.g618950	quick	1	upstream gene variant	MODIFIER
		Cre14.g618900	quick	1	downstream gene variant	MODIFIER
	1,681,521	Cre14.g619100	interm	1	upstream gene variant	MODIFIER
		Cre14.g619166	interm	1	upstream gene variant	MODIFIER
		Cre14.g619133	interm	1	intron variant	MODIFIER
	1,794,162	Cre14.g619800	slow	1	upstream gene variant	MODIFIER
		Cre14.g619850	slow	1	downstream gene variant	MODIFIER
		Cre14.g619825	slow	1	intron variant	MODIFIER
	2,510,754	Cre14.g624950	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	2,917,901	Cre14.g627576	slow	1	intron variant	MODIFIER
	2,925,005	Cre14.g627576	interm	1	intron variant	MODIFIER
	3,218,703	Cre14.g629200	quick	1	upstream gene variant	MODIFIER
		Cre14.g629241	quick	1	downstream gene variant	MODIFIER
		Cre14.g629200-Cre14.g629241	quick	1	intergenic region	MODIFIER
	3,780,281	Cre14.g632501	quick	1	upstream gene variant	MODIFIER
		Cre14.g632400	quick	1	downstream gene variant	MODIFIER
		Cre14.g632450	quick	1	intron variant	MODIFIER
15	298,175	Cre15.g635150	slow	1	conservative inframe insertion	MODERATE
		Cre15.g635200	slow	1	upstream gene variant	MODIFIER
	385,765	Cre15.g640900	interm	1	downstream gene variant	MODIFIER
		Cre15.g640901	interm	1	intron variant	MODIFIER
	434,672	Cre15.g635717	interm	1	downstream gene variant	MODIFIER
		Cre15.g635717	interm	1	downstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre15.g635750	interm	1	intron variant	MODIFIER
		Cre15.g635750	interm	1	intron variant	MODIFIER
	555,528	Cre15.g636450	interm	1	upstream gene variant	MODIFIER
		Cre15.g636350	interm	1	downstream gene variant	MODIFIER
		Cre15.g636400	interm	1	intron variant	MODIFIER
		Cre15.g636400	interm	1	intron variant	MODIFIER
	602,821	Cre15.g636650	interm	1	downstream gene variant	MODIFIER
		Cre15.g636650-Cre15.g636750	interm	1	intergenic region	MODIFIER
	1,204,148	Cre15.g638500	slow	1	upstream gene variant	MODIFIER
		Cre15.g638551	slow	1	intron variant	MODIFIER
	1,341,883	Cre15.g641250	quick	1	5 prime UTR variant	MODIFIER
		Cre15.g641266	quick	1	upstream gene variant	MODIFIER
		Cre15.g641266	quick	1	upstream gene variant	MODIFIER
	1,499,813	Cre15.g642050-Cre19.g750097	quick	1	intergenic region	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	1,547,075	Cre19.g750397	slow	1	upstream gene variant	MODIFIER
		Cre19.g750447	slow	1	upstream gene variant	MODIFIER
		Cre19.g750397-Cre19.g750447	slow	1	intergenic region	MODIFIER
	1,794,425	Cre02.g141626	interm	1	upstream gene variant	MODIFIER
		Cre02.g141666	interm	1	downstream gene variant	MODIFIER
		Cre02.g141666	interm	1	downstream gene variant	MODIFIER
		Cre02.g141646	interm	1	intron variant	MODIFIER
	2,498,470	Cre22.g753997	interm	1	upstream gene variant	MODIFIER
		Cre22.g753947-Cre22.g753997	interm	1	intergenic region	MODIFIER
		Cre22.g753997	slow	1	upstream gene variant	MODIFIER
		Cre22.g753947-Cre22.g753997	slow	1	intergenic region	MODIFIER
	2,583,304	Cre15.g642539-Cre15.g642865	slow	1	intergenic region	MODIFIER
	2,954,314	Cre04.g224931	slow	1	upstream gene variant	MODIFIER
		Cre04.g224931	slow	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre04.g224931	slow	1	upstream gene variant	MODIFIER
		Cre04.g224947	slow	1	downstream gene variant	MODIFIER
		Cre04.g224947-Cre15.g643385	slow	1	intergenic region	MODIFIER
	2,954,330	Cre04.g224931	slow	1	upstream gene variant	MODIFIER
		Cre04.g224931	slow	1	upstream gene variant	MODIFIER
		Cre04.g224931	slow	1	upstream gene variant	MODIFIER
		Cre04.g224947	slow	1	downstream gene variant	MODIFIER
		Cre04.g224947-Cre15.g643385	slow	1	intergenic region	MODIFIER
	3,496,890	Cre20.g751447-Cre23.g755047	interm	1	intergenic region	MODIFIER
	3,796,093	Cre02.g143287-CHR_END	interm	1	intergenic region	MODIFIER
16	272,187	Cre16.g694250	quick	1	upstream gene variant	MODIFIER
		Cre16.g694403	quick	1	downstream gene variant	MODIFIER
		Cre16.g694300	quick	1	intron variant	MODIFIER
		Cre16.g694250	interm	1	upstream gene variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre16.g694403	interm	1	downstream gene variant	MODIFIER
		Cre16.g694300	interm	1	intron variant	MODIFIER
	437,347	Cre16.g693203	slow	1	missense variant	MODERATE
		Cre16.g693202	slow	1	upstream gene variant	MODIFIER
	475,776	Cre16.g692900	slow	1	downstream gene variant	MODIFIER
		Cre16.g692902	slow	1	downstream gene variant	MODIFIER
		Cre16.g692901	slow	1	intron variant	MODIFIER
	493,316	Cre16.g692902	slow	1	upstream gene variant	MODIFIER
		Cre16.g692750	slow	1	upstream gene variant	MODIFIER
		Cre16.g692751	slow	1	upstream gene variant	MODIFIER
		Cre16.g692800	slow	1	intron variant	MODIFIER
	519,168	Cre16.g692585	interm	1	upstream gene variant	MODIFIER
		Cre16.g692550	interm	1	intron variant	MODIFIER
	777,925	Cre16.g647500	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre16.g647500	interm	1	intron variant	MODIFIER
	797,293	Cre16.g647602	slow	1	downstream gene variant	MODIFIER
		Cre16.g647602	slow	1	intron variant	MODIFIER
	801,084	Cre16.g647602	interm	1	frameshift variant	HIGH
		Cre16.g647602	interm	1	frameshift variant	HIGH
	1,450,311	Cre16.g652750	interm	1	downstream gene variant	MODIFIER
		Cre16.g652850	interm	1	downstream gene variant	MODIFIER
		Cre16.g652800	interm	1	intron variant	MODIFIER
	2,044,962	Cre16.g657750	slow	1	upstream gene variant	MODIFIER
		Cre16.g657600	slow	1	downstream gene variant	MODIFIER
		Cre16.g657650	slow	1	downstream gene variant	MODIFIER
		Cre16.g657600	slow	1	downstream gene variant	MODIFIER
		Cre16.g657700	slow	1	intron variant	MODIFIER
	3,239,131	Cre16.g666576	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	3,813,463	Cre16.g689201	slow	1	downstream gene variant	MODIFIER
		Cre16.g689250	slow	1	intron variant	MODIFIER
	4,034,209	Cre16.g670652	interm	1	upstream gene variant	MODIFIER
		Cre16.g670652-Cre16.g670550	interm	1	intergenic region	MODIFIER
	4,382,782	Cre16.g667750	slow	1	3 prime UTR variant	MODIFIER
		Cre16.g667800	slow	1	upstream gene variant	MODIFIER
		Cre16.g667729	slow	1	downstream gene variant	MODIFIER
		Cre16.g667700	slow	1	downstream gene variant	MODIFIER
		Cre16.g667729	slow	1	downstream gene variant	MODIFIER
	4,843,947	Cre16.g686501	interm	1	conservative inframe insertion	MODERATE
		Cre16.g686500	interm	1	upstream gene variant	MODIFIER
	5,078,695	Cre16.g684650	slow	1	upstream gene variant	MODIFIER
		Cre16.g684750	slow	1	downstream gene variant	MODIFIER
		Cre16.g684700	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	5,844,576	Cre16.g678851	interm	1	3 prime UTR variant	MODIFIER
		Cre16.g678850	interm	1	upstream gene variant	MODIFIER
		Cre16.g678750	interm	1	upstream gene variant	MODIFIER
		Cre16.g678700	interm	1	upstream gene variant	MODIFIER
	6,418,147	Cre16.g674450	slow	1	upstream gene variant	MODIFIER
		Cre16.g674500	slow	1	intron variant	MODIFIER
	6,726,301	Cre16.g671850	slow	1	upstream gene variant	MODIFIER
		Cre16.g671750	slow	1	downstream gene variant	MODIFIER
		Cre16.g671800	slow	1	intron variant	MODIFIER
	6,991,246	Cre16.g676533	interm	1	3 prime UTR variant	MODIFIER
		Cre16.g676533	interm	1	3 prime UTR variant	MODIFIER
		Cre16.g676421	interm	1	upstream gene variant	MODIFIER
	7,008,482	Cre16.g676533	interm	1	conservative inframe deletion	MODERATE
		Cre16.g676533	interm	1	conservative inframe deletion	MODERATE

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
		Cre16.g676645	interm	1	upstream gene variant	MODIFIER
		Cre16.g676757	interm	1	downstream gene variant	MODIFIER
		Cre16.g676757	interm	1	downstream gene variant	MODIFIER
	7,370,701	Cre16.g683147	slow	1	3 prime UTR variant	MODIFIER
		Cre16.g683035	slow	1	upstream gene variant	MODIFIER
	7,605,600	Cre16.g687406	interm	1	upstream gene variant	MODIFIER
		Cre16.g687406	interm	1	upstream gene variant	MODIFIER
		Cre16.g687630	interm	1	downstream gene variant	MODIFIER
		Cre16.g687630	interm	1	downstream gene variant	MODIFIER
		Cre16.g687518	interm	1	intron variant	MODIFIER
	7,708,394	Cre16.g689535	slow	1	upstream gene variant	MODIFIER
		Cre16.g689647	slow	1	upstream gene variant	MODIFIER
		Cre16.g689535-Cre16.g689647	slow	1	intergenic region	MODIFIER
17	147,544	Cre17.g697150	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	227,311	Cre17.g697650	interm	1	intron variant	MODIFIER
	235,339	Cre17.g697800	quick	1	upstream gene variant	MODIFIER
		Cre17.g697650	quick	1	downstream gene variant	MODIFIER
		Cre17.g697701	quick	1	downstream gene variant	MODIFIER
		Cre17.g697750	quick	1	downstream gene variant	MODIFIER
		Cre17.g697701-Cre17.g697750	quick	1	intergenic region	MODIFIER
	841,961	Cre17.g702451	slow	1	upstream gene variant	MODIFIER
		Cre17.g702351	slow	1	downstream gene variant	MODIFIER
		Cre17.g702400	slow	1	intron variant	MODIFIER
	999,908	Cre17.g703400	slow	1	3 prime UTR variant	MODIFIER
		Cre17.g703450	slow	1	upstream gene variant	MODIFIER
	1,957,433	Cre17.g710950	slow	1	downstream gene variant	MODIFIER
		Cre17.g711050	slow	1	downstream gene variant	MODIFIER
		Cre17.g711000	slow	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	2,488,437	Cre17.g716150	quick	1	downstream gene variant	MODIFIER
		Cre17.g716101	quick	1	intron variant	MODIFIER
	2,823,518	Cre17.g719325	quick	1	intron variant	MODIFIER
	3,843,005	Cre17.g728000	interm	1	upstream gene variant	MODIFIER
		Cre17.g727950	interm	1	intron variant	MODIFIER
	4,624,735	Cre17.g733050-Cre17.g733100	quick	1	intergenic region	MODIFIER
		Cre17.g733050-Cre17.g733100	interm	1	intergenic region	MODIFIER
	4,891,587	Cre17.g735650	interm	1	missense variant	MODERATE
	4,968,828	Cre17.g736100	slow	1	intron variant	MODIFIER
	5,002,897	Cre17.g736250	slow	1	downstream gene variant	MODIFIER
		Cre17.g736300	slow	1	intron variant	MODIFIER
	5,452,301	Cre17.g739426	interm	1	upstream gene variant	MODIFIER
		Cre17.g739350	interm	1	downstream gene variant	MODIFIER
		Cre17.g739400	interm	1	intron variant	MODIFIER

Chromosome	Position	Gene	Selective history	Nb.replicates	Annotation	Impact score
	5,459,186	Cre17.g739400	interm	1	upstream gene variant	MODIFIER
		Cre17.g739450	interm	1	upstream gene variant	MODIFIER
		Cre17.g739426	interm	1	intron variant	MODIFIER
	5,633,870	Cre17.g740510	interm	1	upstream gene variant	MODIFIER
		Cre17.g740470	interm	1	downstream gene variant	MODIFIER
		Cre17.g740430	interm	1	intron variant	MODIFIER
18	531,627	Cre02.g141126	quick	1	downstream gene variant	MODIFIER
		Cre02.g141106	quick	1	downstream gene variant	MODIFIER
		Cre02.g141106	quick	1	downstream gene variant	MODIFIER
		Cre02.g141106-Cre02.g140941	quick	1	intergenic region	MODIFIER
	1,248,219	Cre17.g733800	slow	1	disruptive inframe insertion	MODERATE

Appendix 7: PDF version of Figure 1 (Molecular function GO terms heatmap for both SR and VR experiments)

See page below

Molecular Functions

zinc ion binding
uridylyltransferase activity
unfolded protein binding
ubiquitin−like modifier activating enzyme activity
transporter activity
transmembrane transporter activity
translation release factor activity, codon specific
tocopherol cyclase activity
sulfiredoxin activity
sulfate transmembrane transporter activity
sodium:dicarboxylate symporter activity
serine−type carboxypeptidase activity
secondary active sulfate transmembrane transporter activity
scavenger receptor activity
quinolinate synthetase A activity
pyridoxal phosphate binding
proton−transporting ATPase activity, rotational mechanism
proton−transporting ATP synthase activity, rotational mechanism
protein tyrosine phosphatase activity
protein serine/threonine phosphatase activity
protein kinase activity
protein binding
protein−disulfide reductase activity
potassium ion transmembrane transporter activity
phosphorus−oxygen lyase activity
peptide−methionine (S)−S−oxide reductase activity
peptidase activity
pantothenate kinase activity
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced ascorbate as one donor, and incorporation of one atom of oxygen
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen
oxidoreductase activity, acting on a sulfur group of donors, disulfide as acceptor
oxidoreductase activity
nucleotide binding
nucleic acid binding
N−acetyltransferase activity
monooxygenase activity
microtubule motor activity
microtubule binding
methyltransferase activity
metal ion binding
iron ion binding
ion channel activity
hydrolase activity, hydrolyzing O−glycosyl compounds
hydrolase activity, acting on acid anhydrides
hydrolase activity
heme binding
GTP binding
endopeptidase inhibitor activity
DNA binding
DNA−binding transcription factor activity
cysteine−type peptidase activity
copper ion binding
chitin binding
cation binding
catalytic activity
calcium ion binding
ATP hydrolysis activity
ATP binding
alpha−amylase activity
acyltransferase activity, transferring groups other than amino−acyl groups
6−phosphofructokinase activity
4 iron, 4 sulfur cluster binding

Counts
12
9
6
3

SR
VR quick
VR interm
VR slow