Australian Government
Cotton Research and Development Corporation

FINAL REPORT 2015

For Public Release

Part 1 - Summary Details

Please use your TAB key to complete Parts 1 & 2.

CRDC Project Number: UQ1301

Project Title: Can genetic diversity predict the potential for emergent glyphosate resistance?

Project Commencement Date: 27/08/2012 **Project Completion Date:** 28/08/2015

CRDC Research Program: 1 Farmers

Part 2 – Contact Details

Administrator: Dr Blake Chapman
Organisation: The University of Queensland
Postal Address: Research Management Office, UQ Research and Innovation, The University of Queensland, Brisbane Queensland 4072, Australia
Ph: +61 7 334 69966 **Fax:** N/A **E-mail:** sciencegrants2@research.uq.edu.au

Principal Researcher: Dr James Hereward, Post Doctoral Researcher
Organisation: The University of Queensland
Postal Address: Goddard Building, The University of Queensland, St. Lucia, Brisbane, 4072
Ph: 07 3365 2755 **Fax:** N/A **E-mail:** j.hereward@uq.edu.au

Supervisor: A. Prof. Gimme Walter
Organisation: The University of Queensland
Postal Address: Goddard Building, The University of Queensland, St. Lucia, Brisbane, 4072
Ph: 07 3365 2228 **Fax:** N/A **E-mail:** g.walter@uq.edu.au

Signature of Research Provider Representative: 

Date Submitted: 12/10/2015

Part 3 – Final Report

(The points below are to be used as a guideline when completing your final report.)

Background

1. Outline the background to the project.

Glyphosate continues to be the world's most important herbicide (Powles & Preston 2006), but in recent years the incidence of resistance has been rising steadily; at least 24 weed species have confirmed resistant populations globally (www.weedscience.org). Resistance to herbicides in general can be broadly categorised as "Target Site" or "Non – Target Site Resistance" the former is largely the result of structural changes to the protein target of the herbicide, the latter broadly includes any other mechanism and its importance has only recently been appreciated (Delye et al 2013). Non-Target site resistance is particularly difficult to decipher as it is usually polygenic and can be constitutive, stress-induced, or possibly both (Delye et al 2013; Cummins et al 2009). NTSR may also confer cross resistance between herbicides of different modes of action, but this remains difficult to predict because so little is known regarding the mechanisms underlying this category of herbicide resistance (Delye 2013).

Resistance to glyphosate is conferred by target site mutations at two amino acid sites in the EPSPS gene, namely Thr^{97} > Ile and Pro^{101} > Thr/Ser (Funke et al 2009) and no other resistant mutations have been established in natural populations of weeds to date. Resistance may also be conferred by EPSPS gene amplification, whereby multiple functional copies of the gene are found throughout the genome of the weed in question (Gaines et al 2010). Non-target site mechanisms are also widespread in glyphosate resistant weeds, in particular the "reduced translocation mechanism", which is more prevalent than target site resistance (Powles & Yu 2010). Currently the reduced translocation mechanism is believed to be the result of a single gene, but the identity of this gene is unknown (Powles & Preston 2006). Several genes have been implicated in NTSR resistance in *Conyza canadensis* via transcriptomic analysis (Peng et al. 2010), and it is likely that estimates of gene number based on F2 backcrosses will be underestimates of the real number of genes underlying translocation mechanisms. Metabolism is another common NTSR mechanism to other herbicides, glyphosate metabolism is known to occur in some weed species via an uncharacterised GOX-like gene pathway, however, metabolism does not result in resistance in the species that have been investigated to date (Sammons & Gaines 2014). It is likely, however, that there are many other NTSR mechanisms that confer glyphosate resistance, but the biochemical and molecular mechanisms have remained elusive to date.

Technological advances have dramatically reduced the cost of sequencing over the past 10 years (Fig. 1). Sequencing cost is likely to drop further during the next few years, but also the speed of sequencing is likely to increase thanks largely to Nanopore sequencing technologies. The advent of the availability of "Omics" approaches to decipher NTSR mechanisms should allow these complex traits to be investigated in ways that were not previously possible (Delye 2013; Delye et al 2013). Indeed, transcriptomics approaches have implicated a number of transporter genes as potential candidates for reduced translocation of glyphosate in Horseweed (*Conyza canadensis*) (Yuan et al 2010; Peng et al 2010). This idea frames the simple question underlying this project; how can we understand glyphosate resistance in cotton system weeds by leveraging new genomic technologies?

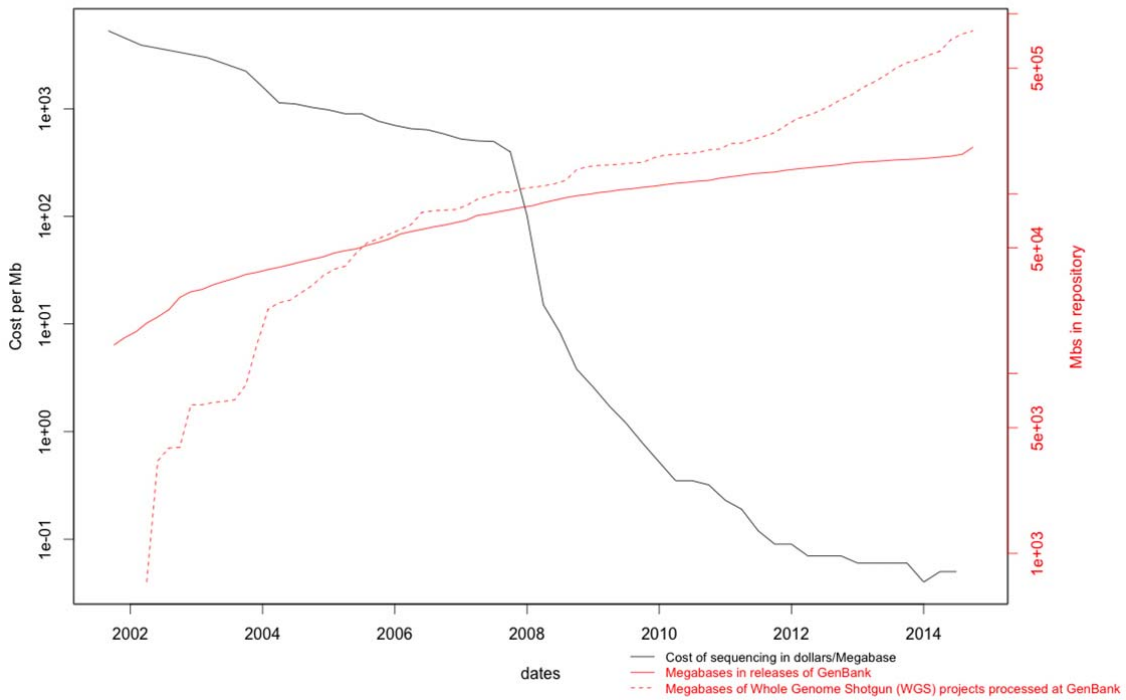

Figure 1. Dramatic falls in DNA sequencing cost since 2002, the inflection point in 2008 represents the introduction of Illumina's sequencing by synthesis approach. The less dramatic rise in genome data on Genbank is due to data reduction during genome assembly and the analysis bottleneck. Image courtesy of Pivotal website.

Objectives

1. To what extent is glyphosate resistance mediated by target site, target amplification, or non-target site mechanisms in four species of resistant weeds in cotton agro ecosystems.
2. How can genomics approaches be used to understand non-target site resistance mechanisms.
3. Can we produce a good draft genome for a weed species in less than $20k?

Methods

1. EPSPS Target site analysis.

Conyza bonariensis

Fleabane has two independent copies of the EPSPS gene; these two copies are highly divergent with approx. 18% DNA bases different between the two (Fig. 2). These two copies are also present in *Conyza canadensis* and *Conyza sumatrensis* meaning that there was a single gene duplication event in the ancestor of these three species, likely millions of years ago, rather than a recent multiple gene duplication event as is the case for Palmer Ameranth in the USA. So that these two copies could be assessed independently I developed two different sets of primers, one that would amplify only EPSPS copy 1 and one that would amplify only EPSPS copy 2. These two primer pairs provide around 200bp of sequence that covers the Proline101 and Threonine97 target site mutations of the EPSPS gene.

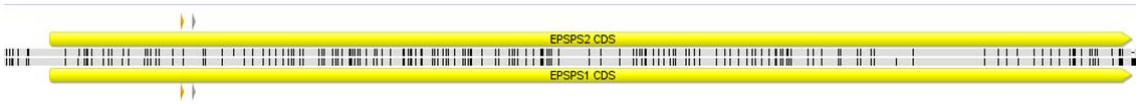

Figure 2. The two copies of the EPSPS gene in *Conyza bonariensis* (fleabane) black lines indicate bases that are different between the two copies, yellow arrows represent the coding sequence of the gene, and the orange and grey arrows indicate the target site mutations.

These primers were then used to amplify the two copies in multiple individuals from four lines, three resistant (Q5, Q42 and Q31) and one susceptible line (Q21). Sanger sequencing was performed by Macrogen (Korea), sequences were edited in codon code aligner, and the clean sequences were aligned against the two EPSPS copies to check for mutations in the target site.

Transcriptome reads generated to assess non-target site resistance in fleabane (see section below) were also mapped against each of the two EPSPS copies, this experiment included slightly different lines to those assessed by sanger sequencing (resistant = Q17, Q42, Q28, susceptible = Q21 and Q34). This method also allows the relative expression level of these genes to be quantified.

Chloris virgata and Chloris truncata

For Windmill grass (*Chloris truncata*) one resistant individual from Tamworth was selected, and for Feathertop Rhodes (*Chloris virgata*) one suspected resistant individual from Aberdeen was used. Both of these samples had DNA extracted, fragmented, and then prepared for Illumina sequencing using the NebNext kit (New England Biolabs). Sequencing was performed on a HiSeq2500 by Novogene (China) generating approximately 20mio Paired-end 125bp reads for each library.

Echinochloa colona

An *E. colona* EPSPS gene was reconstructed by mapping reads to the *Elusine indica* EPSPS gene downloaded from genbank (Accession AJ417033), and then denovo assembly of the *E. colona* reads that matched the gene. The target site resistance status of barnyard grass (*E. colona*) was assessed by mapping transcriptome reads generated for NTSR analysis (see next section) to the *E. colona* EPSPS sequence. This method allows both the nucleotide sequence (target site resistance mutations) and the expression level of EPSPS transcripts to be assessed.

2. E. colona transcriptome experiment

One population of this species in New South Wales was confirmed glyphosate resistant in 2007 (Storrie et al 2008), and there are now multiple resistant populations in Queensland and New South Wales, with resistance also reported in the Ord valley in Northern Western Australia (Gaines et al 2012). I assessed gene expression changes of resistant and susceptible individuals in response to glyphosate treatment with a transcriptomic approach to elucidate the molecular mechanisms for NTSR in this species.

Six lines of *E. colona* were grown for this experiment; the lines were derived from wild populations of barnyard grass originally sampled in Queensland and Northern New South Wales. These lines had been previously screened for resistance by Tony Cook and Jeff Werth, and had undergone one generation in a glasshouse. 24 individuals from each line were raised to at least the two tiller stage in a temperature controlled glasshouse maintained at a maximum temperature of $24^{O}C$, these lines comprised two strongly resistant to glyphosate (PLG3 and QBG41) two with intermediate resistance (PJ55 and QBG3) and two susceptible lines (TC and QB61) so gene expression could be related to the strength of resistance.

Total RNA was extracted from one young growing leaf from each of the 24 individuals of each line 24 hours prior to glyphosate application, glyphosate (without additives) was applied at the standard label rate to all plants, and total RNA was then extracted from a second young growing leaf 24 hours after application. Each individual plant (n = 144) was photographed daily for the following two weeks so that the response to glyphosate treatment could be assessed (Fig. 3). Two individuals from each line were then selected based on their response to glyphosate treatment and used for the subsequent analysis of gene expression changes.

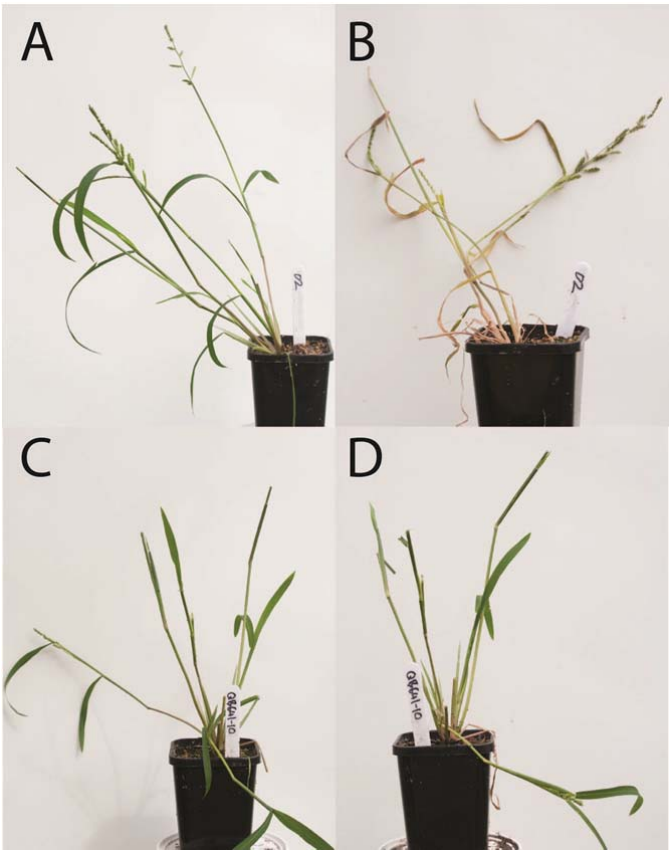

Figure 3. Response of Resistant and Susceptible individuals to glyphosate treatment; A. Susceptible individual from line QB61 on day of treatment, B. same individual one week post treatment, C. Resistant individual from line QBG41 on day of treatment, D. same individual one week post treatment.

The total RNA was subjected to a DNAse treatment to remove any contaminant DNA, and messenger RNA (mRNA) was isolated from the total RNA using a magnetic bead mRNA isolation kit (New England Biolabs). The 24 mRNA samples were then prepared for Illumina sequencing with the NEBNext Ultra Directional RNA kit (New England Biolabs), during this

process DNA “barcodes” or “indexes” are added to each sample so that they can be separated computationally following sequencing. Two pools of samples were prepared, each containing equimolar amounts of each individual library. The first pool comprised two individuals and the pre- and post-glyphosate treatments for the lines TC, PJ55 and QBG41, the second pool comprised all samples for lines QB61, QBG3, and PLG3. Each of these pools was then sequenced on a lane of Illumina HiSeq 2500 by single end 100bp sequencing by Macrogen (Korea).

Raw sequencing reads were checked for quality using FastQC, and adapter contamination removed using Trimmomatic (Lohse et al 2012). A reference transcriptome was assembled using Trinity (Grabherr et al 2011). Reads from each indexed library were then mapped to the reference transcriptome using Bowtie (Langmead et al 2009), allowing two mismatches per read. These alignments were used to create a count table, and differential gene expression was analysed with edgeR (Robinson et al 2010) with a p-value cut-off of 0.05 using the (Benjamini-Hochberg 1995) method for multiple testing correction.

Glyphosate is likely to affect the expression of multiple genes due to its mode of action on a major metabolic gene; a list of all differentially expressed (DE) genes in the two susceptible lines (TC & QB61) was made. These genes were subtracted from the lists of DE genes in all other lines using a custom perl script, so that putative resistance genes could be assessed.

3. *Conyza bonariensis* transcriptome experiment

Five lines of fleabane from Queensland were selected for this experiment, three resistant (Q17, Q42 and Q28) and two susceptible (Q21 and Q34). Seeds were grown under controlled environmental conditions (25^{O}C daily maximum) and artificial lighting. For each line, 9 plants were selected for the experiment (total 49 plants). At approximately the 8-leaf rosette stage total RNA was collected by homogenising a single leaf into TriZol reagent (Life technologies), plants were then immediately sprayed with glyphosate 360 diluted 10ml into one litre. 48hours after glyphosate treatment a second total RNA collection was made. Plants were then monitored for survival (Fig. 4).

Fig. 4, Response of resistant line Q17 to glyphosate over time, although a lot of leaves are lost the survivors are able to regrow from the young growing shoots.

For each line (5 lines) three biological replicates were selected, and for each sample there is a Pre-glyphosate and Post-glyphosate treatment (total RNA samples = 30). Total RNA was subjected to a DNAse treatment to remove any contaminant DNA, and messenger RNA (mRNA) was isolated from the total RNA using a magnetic bead mRNA isolation kit (New England Biolabs). The 30 mRNA samples were then prepared for Illumina sequencing with the NEBNext Ultra Directional RNA kit (New England Biolabs), during this process DNA “barcodes” or “indexes” are added to each sample so that they can be separated computationally following sequencing. All 30 samples were sequenced as a single pool by Novogene (China) across three lanes of a Hiseq2500 producing 125bp paired end reads.

Computational analysis of the data was as described above for *E. colona* briefly, quality control, adapter and quality trimming, denovo assembly of a reference transcriptome, mapping of raw reads to the reference, and differential expression analysis in egdeR.

4. Genome assembly

I have been working on the denovo sequencing and genome assembly of fleabane and barnyard grass. For each species two individuals of susceptible and resistant lines have been sequenced at relatively low coverage (approx. 10x) using Illumina Hiseq2500 (paired end 125bp reads). For fleabane resistant lines are Q42, Q31 and Q5, susceptible lines N15 and Q34. For barnyard grass resistant lines are PLG3 and QBG41, intermediate are PJ55 and QBG3, susceptible lines are TC and QBG1. One individual of each species has been sequenced at higher coverage (approx. 100x) for genome assembly. Initial assemblies have been produced using the SOAPdenovo, velvet, and SpAdes assembly algorithms.

Results

1. EPSPS Target site analysis.

Conyza bonariensis

None of the three resistant lines screened by Sanger sequencing (Q5, Q31, Q42) had any resistance-causing mutations at either of the two amino acid positions known to confer resistance in wild weed populations (Table 1). Further, the three resistant lines assessed by transcriptome reads (Q17, Q42, Q28) also showed no evidence of target site resistance. In total five resistant lines were screened and in all of these lines resistance was not caused by target site mutations. Fleabane in Australia therefore must have non-target site mechanisms conferring resistance to glyphosate.

Table 1. Resistance mutation screening of three resistant and one susceptible line of fleabane for both known resistance mutations in both copies of the EPSPS gene.

Individual	Phenotype	EPSPS1 THR97	EPSPS1 PRO101	EPSPS2 THR 97	EPSPS2 PRO101
Q5-11	Resistant	SUS	SUS	SUS	SUS
Q5-10	Resistant	SUS	SUS	SUS	SUS
Q5-09	Resistant	SUS	SUS	NA	NA
Q5-08	Resistant	SUS	SUS	SUS	SUS
Q5-07	Resistant	SUS	SUS	SUS	SUS
Q5-06	Resistant	SUS	SUS	SUS	SUS
Q5-05	Resistant	SUS	SUS	SUS	SUS
Q5-04	Resistant	SUS	SUS	SUS	SUS
Q5-02	Resistant	SUS	SUS	SUS	SUS
Q5-01	Resistant	SUS	SUS	SUS	SUS
Q42-05	Resistant	NA	NA	SUS	SUS
Q42-04	Resistant	SUS	SUS	SUS	SUS
Q42-03	Resistant	SUS	SUS	SUS	SUS
Q42-02	Resistant	SUS	SUS	SUS	SUS
Q42-01	Resistant	SUS	SUS	NA	NA
Q31-11	Resistant	SUS	SUS	SUS	SUS

Q31-10	Resistant	SUS	SUS	SUS	SUS
Q31-09	Resistant	SUS	SUS	SUS	SUS
Q31-08	Resistant	SUS	SUS	SUS	SUS
Q31-07	Resistant	SUS	SUS	SUS	SUS
Q31-06	Resistant	SUS	SUS	SUS	SUS
Q31-05	Resistant	SUS	SUS	NA	NA
Q31-04	Resistant	SUS	SUS	SUS	SUS
Q31-03	Resistant	SUS	SUS	SUS	SUS
Q31-02	Resistant	SUS	SUS	SUS	SUS
Q31-01	Resistant	SUS	SUS	NA	NA
Q21-24	Susceptible	SUS	SUS	SUS	SUS
Q21-23	Susceptible	SUS	SUS	SUS	SUS
Q21-22	Susceptible	SUS	SUS	SUS	SUS
Q21-21	Susceptible	SUS	SUS	SUS	SUS
Q21-20	Susceptible	SUS	SUS	NA	NA
Q21-19	Susceptible	SUS	SUS	SUS	SUS
Q21-18	Susceptible	SUS	SUS	SUS	SUS
Q21-16	Susceptible	SUS	SUS	SUS	SUS
Q21-14	Susceptible	SUS	SUS	SUS	SUS
Q21-13	Susceptible	SUS	SUS	SUS	SUS
Q21-12	Susceptible	SUS	SUS	SUS	SUS
Q21-10	Susceptible	SUS	SUS	SUS	SUS
Q21-09	Susceptible	SUS	SUS	SUS	SUS
Q21-08	Susceptible	SUS	SUS	SUS	SUS
Q21-07	Susceptible	SUS	SUS	SUS	SUS
Q21-06	Susceptible	SUS	SUS	SUS	SUS
Q21-05	Susceptible	SUS	SUS	SUS	SUS
Q21-04	Susceptible	SUS	SUS	SUS	SUS

Previous studies in Spain have found target site resistance mutations in this species, and also have reported higher EPSPS gene expression in resistant lines. However, my analysis of the transcriptome data indicates the same relative expression of EPSPS in resistant and susceptible lines, and also before and after glyphosate treatment. If gene copy number or relative transcription level were a factor in endowing resistance in Australian fleabane then it would be expected that either the resistant lines would have higher expression levels (gene copy number), or that expression in resistant individuals would be up-regulated following glyphosate exposure (transcription response). These results therefore indicate that glyphosate resistance in Australian fleabane has nothing to do with target site mutations or EPSPS expression levels.

Chloris virgata and Chloris truncata

Windmill grass (*Chloris truncata*), shows the susceptible genotype (no target site mutations) at the EPSPS gene (Fig. 4). This population was confirmed resistant, so it is likely that glyphosate resistance in this population is via NTSR mechanisms. Further sampling and sequencing of this species is probably justified to rule out the presence of target site mutations in other populations. My preliminary analysis also indicates a higher number of reads mapping to the EPSPS gene (approx. 120x coverage) compared to other nuclear genes (approx. 30x), this may be evidence of EPSPS gene amplification, and would indicate 4

copies of EPSPS in the genome. Further work is required to establish whether this is a real result and not an artefact of the choice of nuclear genes; I need to assemble and check more nuclear genes to confirm this result. There are only two haplotypes present in the reads mapping to EPSPS, so if gene amplification has occurred in this species it looks like it was fairly recent.

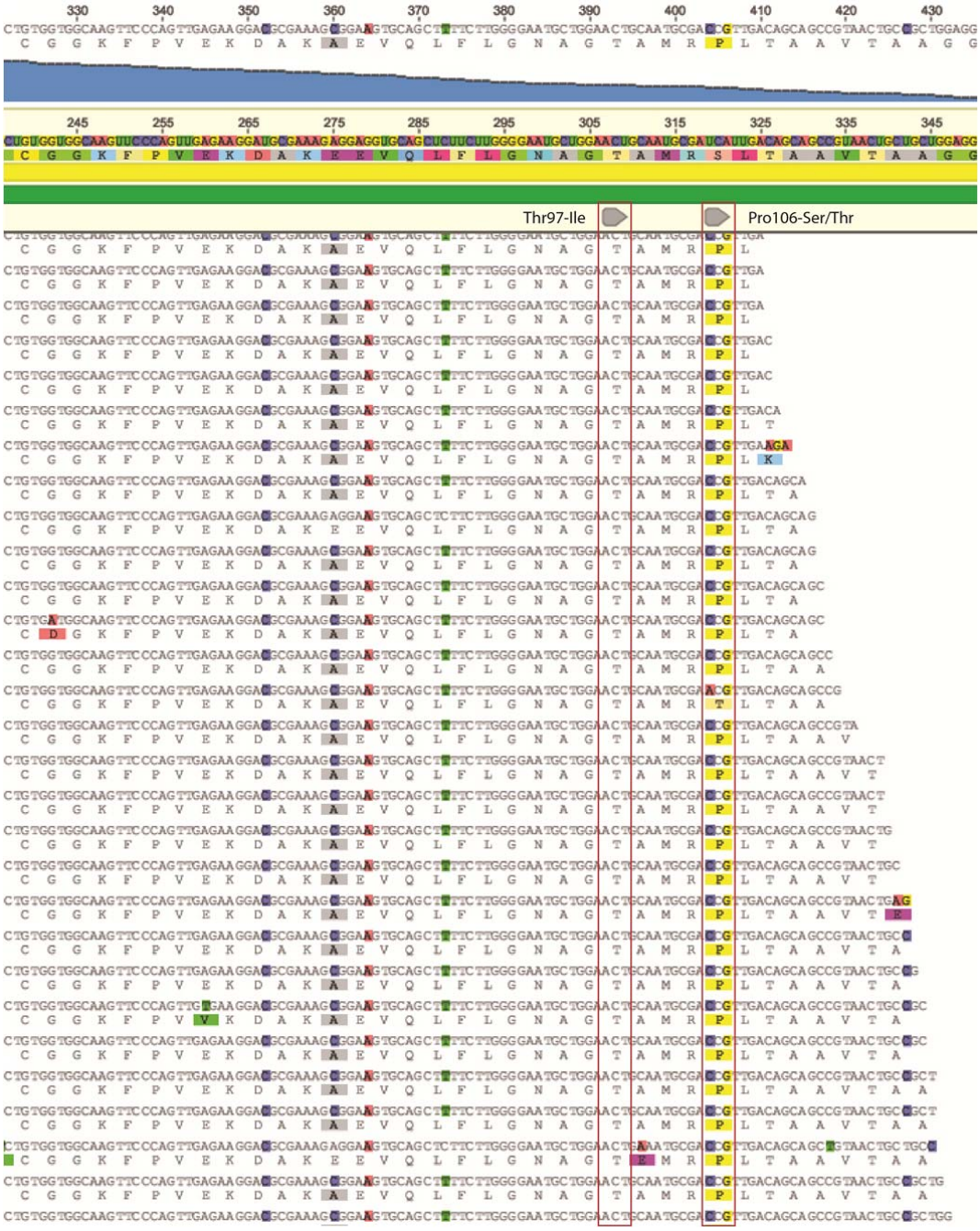

Figure 4. *Chloris truncata* reads mapping to the EPSPS gene, both the Pro106 and Thr97 amino acids have the susceptible genotype.

Although the Feathertop Rhodes (*Chloris virgata*) individual was only recorded as suspected resistant, the genotype of this individual has the Proline106 to Serine mutation (the most common glyphosate target site mutation) (Fig 5.). Feathertop Rhodes is not on the glyphosate label and has been considered tolerant to glyphosate, but it has been reported to be surviving higher doses of glyphosate more recently. It is possible that Feathertop had the target site mutation prior to the use of glyphosate, which would explain its initial tolerance. This

scenario would not explain the recent rise in survivorship at higher rates of glyphosate. It is more likely that this weed was naturally tolerant to a certain extent initially, and that that subsequent exposure to glyphosate has resulted in the selection of target site resistance.

The genome sequence indicates a similar number of reads mapping to EPSPS as to other nuclear genes in this species (approx. 20x coverage), so there is no evidence of target site gene copy amplification in this species.

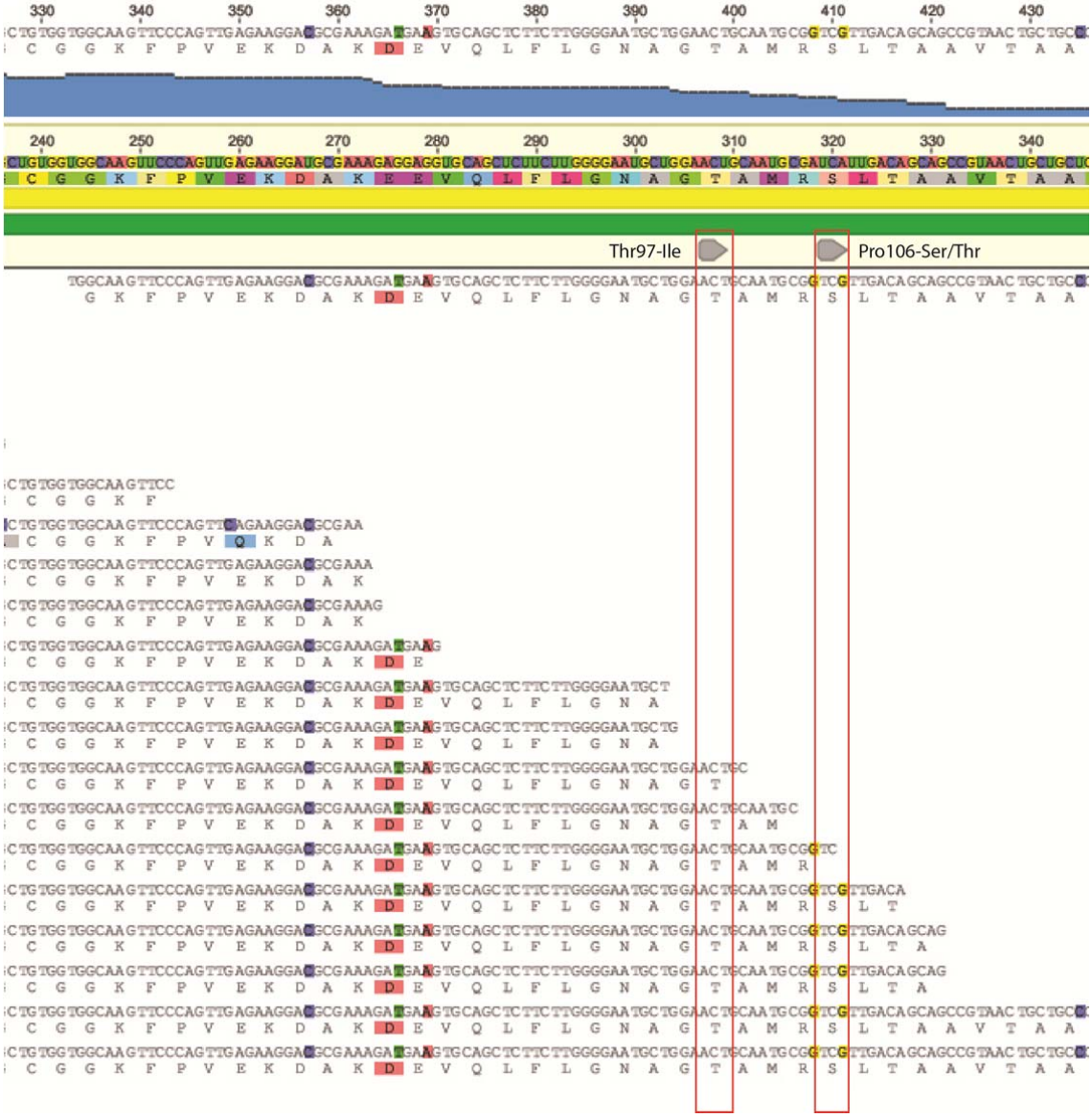

Figure 5. Mapping *Chloris virgata* to the EPSPS gene, showing the Pro106-Serine mutation caused by a mutation to G in the third codon position (right red box).

Echinochloa colona

Coverage of the EPSPS gene was consistent between samples (mean = 44x) when the total number of reads in each library was controlled for, indicating that multiple copies of this

gene are not being expressed in this species. The assessment of EPSPS mutations showed that all individuals are susceptible at the target site with the exception of the strong resistant QBG41, which was heterozygous resistant at the EPSPS Pro^{101} site with both the Proline and Threonine amino acid substitutions being expressed (Fig. 6, Table 2).

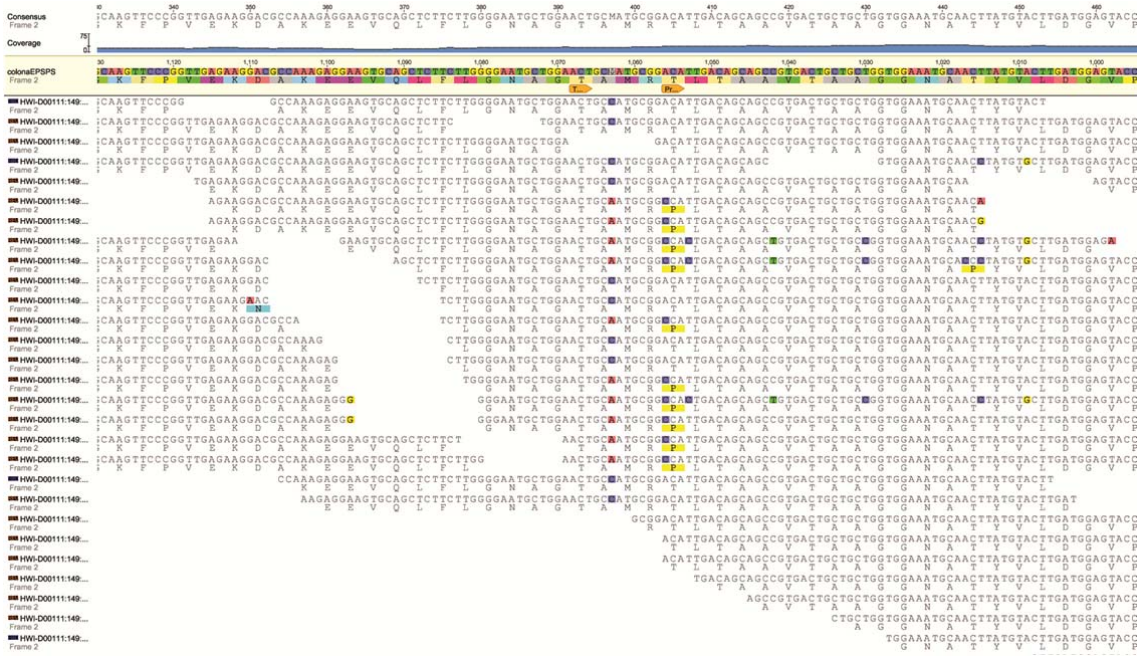

Figure 6. Mapping of QBG41 to *E. colona* EPSPS showing both the Proline (susceptible) and Threonine (resistant) amino acid translations at Proline 106, this plant is heterozygous resistant at the genotype level.

Table 2. Mutations present at the EPSPS target site gene for each of the 12 individual plants sequenced, the two mutations that confer glyphosate resistance were scored and all lines were susceptible at the target site with the exception of QBG41.

SAMPLE	Resistance Phenotype	Pro101	RES	Thr97	RES
TC-05	Susceptible	P	SUS	T	SUS
TC-10	Susceptible	P	SUS	T	SUS
QB61-02	Susceptible	P	SUS	T	SUS
QB61-03	Susceptible	P	SUS	T	SUS
PJ55-04	Intermediate	P	SUS	T	SUS
PJ55-06	Intermediate	P	SUS	T	SUS
QBG3-01	Intermediate	P	SUS	T	SUS
QBG3-02	Intermediate	P	SUS	T	SUS
QBG41-04	Strong	P/T	SUS/RES	T	SUS
QBG41-10	Strong	P/T	SUS/RES	T	SUS
PLG3-09	Strong	P	SUS	T	SUS
PLG3-11	Strong	P	SUS	T	SUS

2. *E. colona transcriptome experiment*

The response of each line to glyphosate confirmed the resistance status that had been previously assessed (Cook, T. A. and Werth J. A. *pers comm.*). The two Illumina sequencing lanes produced a high number of reads for each library (Table 3. mean = 12,000,000), providing adequate coverage to detect differentially expressed genes where expression is not very low.

Table 3. Details of the resistance phenotype as determined by bioassay, glyphosate treatment (24 hours Pre-spray or 24 hours Post-spray) and the number of 100bp sequence reads obtained for each of 24 indexed libraries.

Sample	Resistance Phenotype	Treatment	Reads
TC-05_PRE	Susceptible	PRE	8,555,681
TC-10_PRE	Susceptible	PRE	13,397,815
QB61-02_PRE	Susceptible	PRE	13,072,398
QB61-03_PRE	Susceptible	PRE	15,620,354
PJ55-04_PRE	Intermediate	PRE	12,026,933
PJ55-06_PRE	Intermediate	PRE	8,803,624
QBG3-01_PRE	Intermediate	PRE	9,439,295
QBG3-02_PRE	Intermediate	PRE	12,903,868
QBG41-04_PRE	Strong	PRE	10,684,229
QBG41-10_PRE	Strong	PRE	11,587,148
PLG3-09_PRE	Strong	PRE	14,797,291
PLG3-11_PRE	Strong	PRE	14,257,310
TC-05_POST	Susceptible	POST	10,359,500
TC-10_POST	Susceptible	POST	13,820,611
QB61-02_POST	Susceptible	POST	13,166,377
QB61-03_POST	Susceptible	POST	12,766,913
PJ55-04_POST	Intermediate	POST	11,552,921
PJ55-06_POST	Intermediate	POST	14,837,369
QBG3-01_POST	Intermediate	POST	12,201,423
QBG3-02_POST	Intermediate	POST	11,829,282
QBG41-04_POST	Strong	POST	11,238,106
QBG41-10_POST	Strong	POST	12,517,083
PLG3-09_POST	Strong	POST	9,032,924
PLG3-11_POST	Strong	POST	9,757,971

Each of the treatments had numerous genes that were differentially expressed in response to glyphosate treatment (e.g. Fig. 7), 3,584 of these showed DE in the susceptible lines, and their expression is thus altered due to the physiological response of the plant to glyphosate treatment, and likely not involved in resistance. Once these genes had been removed from the list of DE genes for each of the resistant lines, the two strong resistant lines (PLG3 & QBG41) had a much higher number of DE genes than the two lines with intermediate resistance (Table 4.)

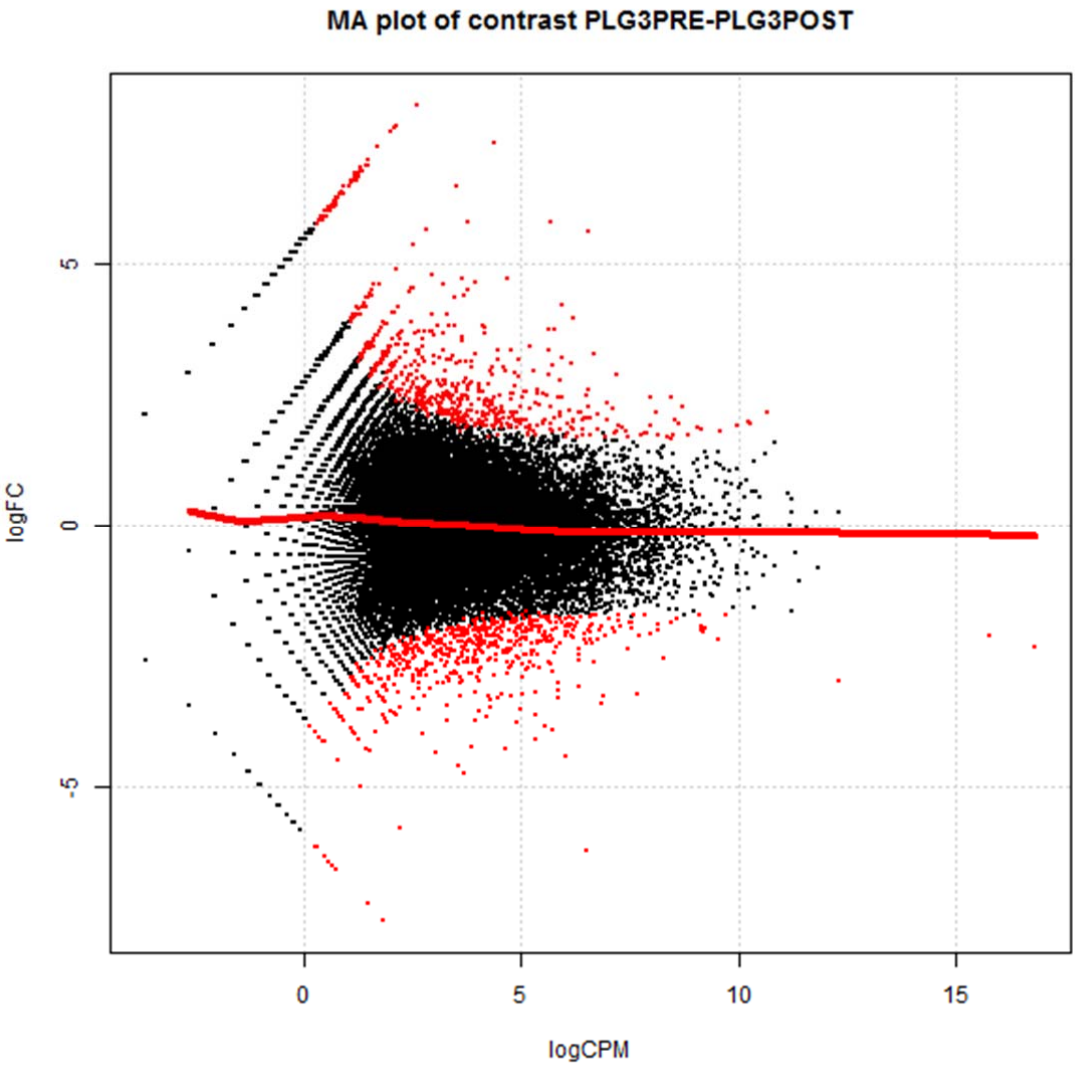

Figure 7. MA plot showing the relationship between log2 fold change and expression strength for the PLG3 line, red dots indicate genes that are significantly differentially expressed.

Table 4. Number of Differentially Expressed genes in each of the resistant lines

Line	Resistance	Number of DE genes
QBG3	Intermediate	44
PJ55	Intermediate	56
PLG3	Strong	267
QBG41	Strong	320

This demonstrates that Non-target site resistance is indeed prevalent in Australian *E. colona*, with only one of the four resistant lines having a target site mutation and no evidence for gene amplification in the lines investigated here. NTSR resistance in *E. colona* is likely polygenic, as indicated by the numerous genes that are differentially expressed in resistant lines, particularly those with strong resistance.

3. *Conyza bonariensis* transcriptome experiment

Line Q28 was recorded as resistant; however, all of the plants were killed by glyphosate application at the label rate, this line is therefore considered as intermediate for the purposes of this analysis. All samples returned more than 20mio reads after quality control and trimming (Table 5.)

Table 5. Summary of data generated, sample name, resistance phenotype and treatment (Pre or Post glyphosate application) and the number of paired end 125bp reads generated for each sample

Sample	Resistance Phenotype	Treatment	Paired end Reads
Q34_S1	Susceptible	PRE	36,873,565
Q34_S2	Susceptible	PRE	29,166,701
Q34_S5	Susceptible	PRE	25,340,474
Q21_S4	Susceptible	PRE	26,685,084
Q21_S5	Susceptible	PRE	27,596,528
Q21_S7	Susceptible	PRE	24,905,681
Q28_R5	Intermediate	PRE	26,216,964
Q28_R6	Intermediate	PRE	23,744,008
Q28_R8	Intermediate	PRE	25,554,635
Q42_R4	Resistant	PRE	25,842,002
Q42_R5	Resistant	PRE	27,667,172
Q42_R6	Resistant	PRE	25,136,147
Q17_R6	Resistant	PRE	23,661,521
Q17_R1	Resistant	PRE	24,903,869
Q17_R5	Resistant	PRE	28,187,739
Q34_S1	Susceptible	POST	29,623,317
Q34_S2	Susceptible	POST	29,290,698
Q34_S5	Susceptible	POST	26,966,113
Q21_S4	Susceptible	POST	29,812,118
Q21_S5	Susceptible	POST	25,931,912
Q21_S7	Susceptible	POST	25,949,531
Q28_R5	Intermediate	POST	25,673,459
Q28_R6	Intermediate	POST	29,804,114
Q28_R8	Intermediate	POST	23,696,055
Q42_R4	Resistant	POST	31,895,529
Q42_R5	Resistant	POST	28,891,687
Q42_R6	Resistant	POST	23,464,420
Q17_R6	Resistant	POST	25,108,396
Q17_R1	Resistant	POST	26,439,492
Q17_R5	Resistant	POST	27,742,193

A very high number of genes were differentially expressed (DE) in fleabane following glyphosate treatment. The susceptible lines had significantly more DE genes than the resistant lines, as they are undergoing the full effects of glyphosate on the functioning of normal enzymatic pathways (Fig. 8). The two resistant lines Q17 and Q42 had 7,121 and 7,846 DE genes respectively, but the two susceptible, Q34 and Q28 had 23,160 and 13,035 DE genes after removing genes with a p value above 0.05 and a fold change below 2.

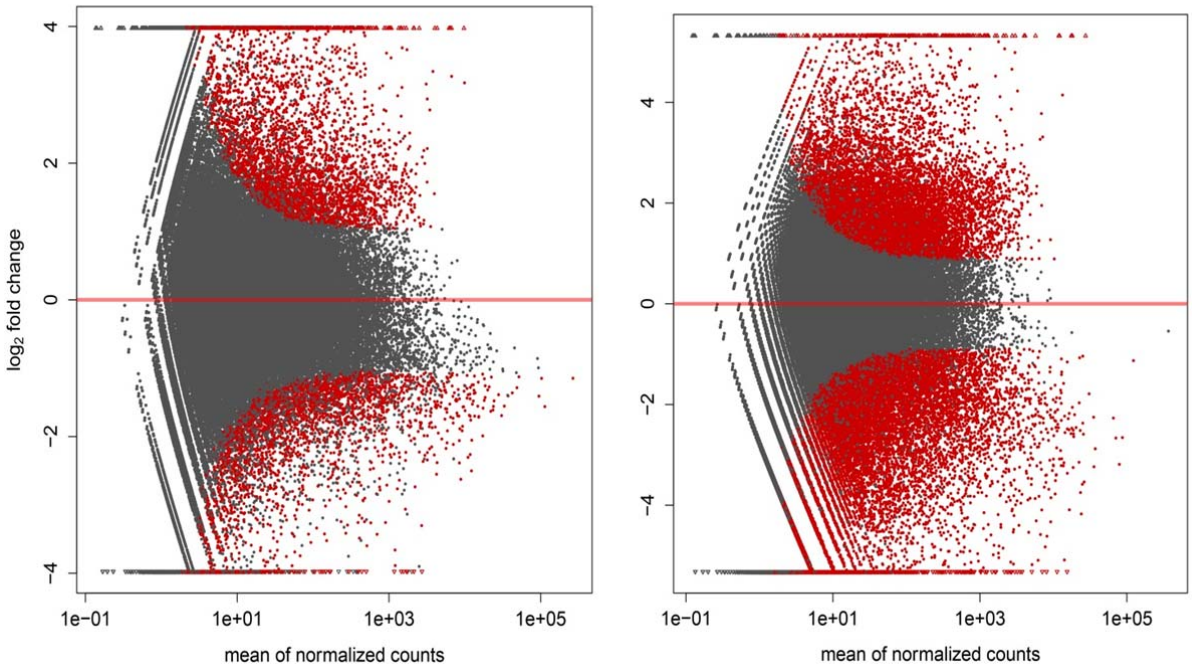

Figure 8. MA plots showing the number of significantly differentially expressed genes (red dots) for Q17 Resistant (left) and Q34 Susceptible (right).

A list of all genes that are differentially expressed in the susceptible lines was created, and then these genes were subtracted from the lists of DE genes in the resistant lines. Following the removal of genes that are differentially expressed in response to glyphosate, line Q17 only had three DE genes remaining, and line Q42 only had two. Blastx searches against known proteins indicated that these genes are involved in plant signalling pathways. The low number of genes that are DE in fleabane may indicate that NTSR in fleabane is enabled by only a few genetic changes. The cut off values used in this analysis may be somewhat crude though, and further analysis of the relative expression changes amongst the most highly expressed genes is warranted.

4. Genome assembly

I have produced genome assemblies have with a total assembly size close to the genome size of each of the species, and the complete chloroplast sequence has been assembled for each species has been successfully reconstructed and annotated (e.g. Fig. 9). These assemblies are, however, heavily fragmented because they have been produced with short read sequencing technology. For the purposes of this project I had allowed $20k per species for genome assembly, which is significantly lower than the cost of previous genomes. I had planned on using PacBio long read single molecule sequencing to improve the assemblies, but the cost and DNA input requirements ended up being too high for this project. Instead, I applied to be a beta tester for the new Oxford Nanopore DNA sequencing device, but have only just received the device after two years of applying to get one (Fig. 10).

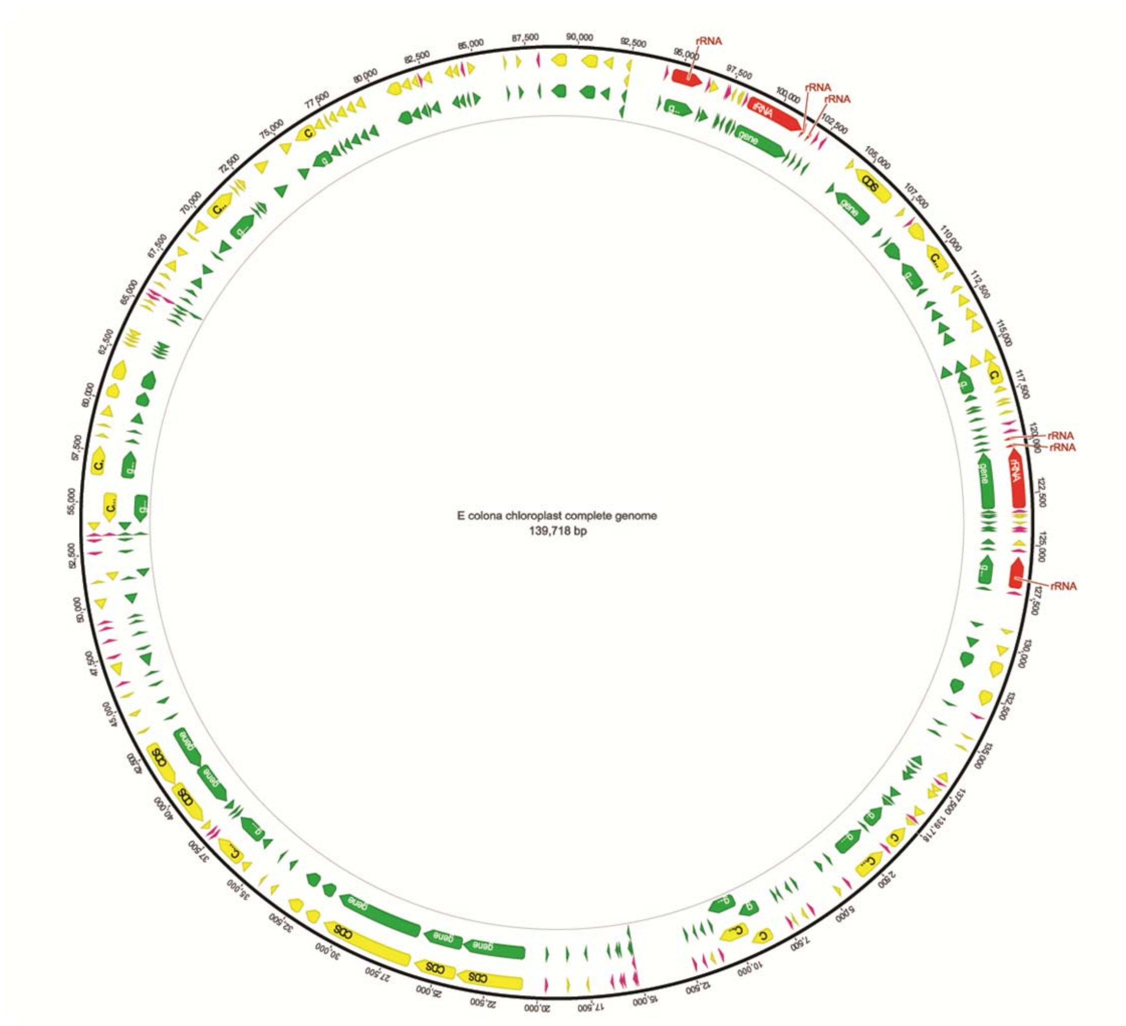

Fig. 9. Complete annotated chloroplast genome for *Echinochloa colona* (139,781bp).

The combination of Nanopore long reads and Illumina high depth and accuracy should allow the full reconstruction of fleabane's genome, which will be an essential tool to unravel NTSR in this species. There are other methods available to generate complete genome assemblies,

however, none of these can currently be accomplished at the price point that I allocated for this part of the research.

My work on barnyard grass has identified that it is an allohexaploid, this means that it has three copies of the same diploid genome, and that each cell therefore carries six genomes.

Fig. 10. Oxford Nanopore sequencer getting ready to sequence the fleabane genome.

Outcomes

We now have a good understanding of the target site status in four important glyphosate resistant weeds; the resistance of only one of these can be explained by target site mutations alone. This highlights the prevalence of NTSR mechanisms in the cotton agro-ecosystem. The industry outcome of this research will be the risk-assessment of the chance of cross resistance between herbicide groups, however, this risk cannot be fully assessed until the underlying mechanisms have been identified. To this end I have made a good start with both barnyard grass and fleabane, producing lists of candidate genes from the transcriptome analyses. I have also been amassing genomic data and tools that will enable NTSR mechanisms to be more fully nailed down in project UQ1501.

UQ1301 also represents CRDC's investment in training and capacity building. I have undertaken several advanced training courses in bioinformatics, as well as becoming proficient in the "wet" and analytical skills required to undertake NGS (Next generation sequencing) experiments. I have also been able to then pass these skills onto other cotton researchers (Sharon Van Brunschot, Murray Sharman, Dean Brookes). I have been furthering my personal development through the co-supervision of three cotton PhD students (Rehan Silva, cruiser fund, Dean Brookes, Justin Cappadonna) and one honours student who is working within Justin's project at the moment. I think the model that we have in the lab currently of several cotton researchers (post doc and PhD) in the one lab group has been particularly successful in terms of maximising the investment in training and capacity building.

Extension Opportunities

2. Detail a plan for the activities or other steps that may be taken:

An article in spotlight highlighting the prevalence of glyphosate NTSR in the cotton system and what it may mean for cross resistance is warranted. Obviously this is a fairly difficult topic to explain because knowledge of NTSR mechanisms is still limited. I think that it is a good time to be explaining why it is worth investing in understanding what NTSR mechanisms are present and what risk of cross resistance they may pose. I have also been asked by Mark Congreve (ICAN) for an extension piece on resistance mechanisms.

9. A. List the publications arising from the research project and/or a publication plan.
I have the following publications planned:

1. Investigation of non-target site mechanisms in fleabane via transcriptome sequencing.
2. Investigation of non-target site mechanisms in barnyard grass via transcriptome sequencing.
3. Target site resistance in Feathertop Rhodes grass (this will be in collaboration with Jeff Werth in the new project).
4. Gene amplification in Windmill grass (following more work).

I am also currently thinking about a review paper on how we can use new genome editing techniques to understand NTRS in weeds.

B. Have you developed any online resources and what is the website address?

I have put some of our lab protocols online at http://www.jameshereward.org/protocols.html including a cheap method of producing high quality DNA by silica spin columns. This could potentially be useful to other CRDC funded genetics researchers. This is something that I have done in my spare time rather than being part of the project.

References

Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1): 289-300.

Cummins, I., D. Bryant and R. Edwards (2009). "Safener responsiveness and multiple herbicide resistance in the weed black-grass (Alopecurus myosuroides)." *Plant biotechnology journal* **7**(8): 807-820.

Délye, C. (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 management science* **69**(2): 176-187.

Délye, C., M. Jasieniuk and V. Le Corre (2013). "Deciphering the evolution of herbicide resistance in weeds." *Trends in genetics : TIG* **29**(11): 649-658.

Funke, T., Y. Yang, H. Han, M. Healy-Fried, S. Olesen, A. Becker and E. Schönbrunn (2009). "Structural basis of glyphosate resistance resulting from the double mutation Thr97 -> Ile and Pro101 -> Ser in 5-enolpyruvylshikimate-3-phosphate synthase from Escherichia coli." *The Journal of biological chemistry* **284**(15): 9854-9860.

Gaines, T., W. Zhang, D. Wang, B. Bukun, S. Chisholm, D. Shaner, S. Nissen, W. Patzoldt, P. Tranel, A. Culpepper, T. Grey, T. Webster, W. Vencill, R. Sammons, J. Jiang, C. Preston, J. Leach and P. Westra (2010). "Gene amplification confers glyphosate resistance in Amaranthus palmeri." *Proceedings of the National Academy of Sciences of the United States of America* **107**(3): 1029-1034.

Grabherr, M., B. Haas, M. Yassour, J. Levin, D. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." *Nature biotechnology* **29**(7): 644-652.

Joshua, S. Y., L. G. A. Laura, C. Yongwei, D. H. Matthew, Z. Xin, P. Yanhui, H. Jun, R. R. Murali, R. H. Gregory, J. L. Thomas, R. D. Sammons, W. Xinwang, R. Priya, H. J. Denita, A. W. Phillip, E. S. Brian, A. R. Timothy, N. T. Robert and C. N. Stewart (2010). "Functional Genomics Analysis of Horseweed (Conyza canadensis) with Special Reference to the Evolution of Non–Target-Site Glyphosate Resistance." *Weed Science* **58**.

Langmead, B., C. Trapnell, M. Pop and S. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." *Genome biology* **10**(3).

Lohse, M., A. Bolger, A. Nagel, A. Fernie, J. Lunn, M. Stitt and B. Usadel (2012). "RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics." *Nucleic acids research* **40**(Web Server issue): 7.

Peng, Y., L. Abercrombie, J. Yuan, C. Riggins, R. Sammons, P. Tranel and C. Stewart (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 management science* **66**(10): 1053-1062.

Powles, S. and Q. Yu (2010). "Evolution in action: plants resistant to herbicides." *Annual review of plant biology* **61**: 317-347.

Powles, S. B. and Preston, C. (2006). "Evolved Glyphosate Resistance in Plants: Biochemical and Genetic Basis of Resistance 1." *Weed Technology* **20**.

Robinson, M., D. McCarthy and G. Smyth (2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics (Oxford, England)* **26**(1): 139-140.

Sammons, R.D. & T. A. Gaines, 2014, “Glyphosate resistance: state of knowledge”, Pest Management Science, 70: 1367–1377.

Storrie, A., T. Cook, P. Boutsalis, D. Penberthy, P. Moylan and R. van Klinken (2008). "Glyphosate resistance in awnless barnyard grass (Echinochloa colona (L.) Link) and its implications for Australian farming systems."

Todd, A. G., C. Andrew and B. P. Stephen (2012). "Evolved Resistance to Glyphosate in Junglerice (Echinochloa colona) from the Tropical Ord River Region in Australia." *Weed Technology* **26**.

Part 4 – Final Report Executive Summary

Resistance to herbicides can happen by changes to the target gene (target site resistance) or other genetic changes (non-target site resistance, NTSR). Cross resistance between different herbicide groups is impossible with target site resistance; however, some NTSR mechanisms may allow a weed to be resistant to several herbicide groups after having only been exposed to one, but evidence to date is scant on this. I therefore examined four weed species from the northern region cotton growing areas and demonstrated that fleabane, barnyard grass and windmill grass all have glyphosate resistance by NTSR mechanisms. The only species investigated that had a fixed target site mutation is Feathertop Rhodes grass. The risk of Feathertop developing cross resistance is therefore low, but in the other three species we need to understand the mechanism of NTSR to assess this risk.

I have used gene expression analysis to test for genes that are more highly produced or supressed in resistant barnyard grass and fleabane to explore NTSR mechanisms in these two species. In barnyard grass about 50 genes are differentially expressed in intermediate resistant lines, whereas about 300 are in strong resistant lines. This indicates that NTSR is a polygenic trait in this weed. In fleabane only five genes were differentially expressed in the resistant lines, and these genes are involved in plant signalling pathways and transporter proteins. It is therefore likely that NTSR in fleabane is controlled by relatively few genes and involves the transport of glyphosate within the plant.

I have developed considerable genomic resources for both fleabane and barnyard grass, and these genomic tools provide the basis for the further elucidation of NTSR mechanisms in these two species. NTSR is still difficult to dissect mechanistically, but the advent of cheap genome sequencing means that it is now more possible than it has been previously. The genomic data generated during this project (UQ1301) will be leveraged in project UQ1501 to establish the mechanisms of NTSR in cotton-system weeds so that the threat of cross resistance can be assessed as the industry moves to stacked herbicide resistance traits.