

# **Novel mutants optimised for lignin, growth, and biofuel production via re-mutagenesis and co-expression analysis**

## **Investigators**

In the Division of Plant Sciences, College of Life Sciences, University of Dundee:

Claire Halpin, Professor and Deputy Head of Division

Gordon Simpson, Lecturer

Christopher McClellan, Postdoctoral researcher

Katarzyna Rataj, Postdoctoral researcher

Yuguo Xiao, Postdoctoral researcher

Abdellah Barakate, Postdoctoral researcher

Carinna Baird, Technician

Sean Hackett, Technician

Natalie Walker, Technician

Lydia Welsh, Undergraduate Student

## **External Collaborators:**

Wout Boerjan, Professor, University of Ghent, Belgium

Simon McQueen-Mason, Professor, University of York, UK

## **Abstract**

The objectives of this project are to find novel genes that will be useful for optimising plant biomass to facilitate biofuel production. These include genes influencing saccharification (the release of sugars from biomass), genes involved in the synthesis or control of lignin, and genes that restore normal growth to lignin mutants. Several approaches are being used towards this goal. Firstly, a founder mutant which has reduced lignin and improved saccharification but stunted growth, was re-mutagenised and 43 secondary mutants have been identified that rescue the reduced size to normal or near-normal growth. While many of these mutants have restored expression of the founder gene, a large portion of these restored mutants still contain significant reductions in lignin and a corresponding increase in sugar release upon saccharification. Six mutants have been selected for mutation mapping to identify the causal mutation. A second type of screen has been performed to identify mutants that enhance sugar release over and above the improvement already achieved in a second founder lignin mutant. This screen is approximately 50% completed with many samples currently being analysed. After re-screening mutants initially identified as having increased saccharification, one mutant has been isolated that significantly increases the release of sugar above that of the founder mutant. Lignin staining of this mutant suggests a collapsed xylem phenotype and decreased lignin levels. Another approach used to identify novel genes involved in lignin synthesis and regulation is co-expression analysis. Over 65 mutants in genes identified as being tightly co-expressed with lignin biosynthesis genes have been analysed. Two of these mutants have been shown to contain reduced lignin and elevated sugar release in the saccharification assay. One novel gene is being studied in detail and a patent is being filed. Protein complex purification and identification has also been used to identify interacting partners of lignin biosynthesis proteins. Two known lignin biosynthesis enzymes and one protein identified from the co-expression analysis have been used as baits for tagged purification strategies. Initial results from LC/MS of proteins interacting with the novel protein identified from co-expression analysis indicates that this protein interacts with multiple lignin biosynthesis enzymes, further supporting a

role for this novel protein in lignin biosynthesis. Continuing work involves mapping mutants identified from the mutagenesis screens and identifying interacting protein partners for several monolignol biosynthesis enzymes.

## **Introduction**

The aim of this project is to identify novel genes involved in the lignin biosynthetic process by screening for plants that have improved saccharification properties or improved growth characteristics. Novel genes are being identified through mutagenesis of existing lignin-defective mutants, as well as screening plants defective in genes that are co-expressed with known lignin biosynthetic genes. Further analysis of lignin-related proteins is being undertaken by isolating interaction partners of known monolignol biosynthesis enzymes and these interaction partners are being identified through liquid chromatography/mass spectroscopy analysis. Genes identified by these strategies will be good candidates for manipulation in crops used in biofuel production.

## **Background**

Although the enzymes involved in producing monolignols in *Arabidopsis thaliana* have been identified, there are still many questions about how lignin synthesis is regulated and controlled. Recently, it has been found that control of the synthesis of upstream metabolites, such as phenylalanine has an effect on lignin levels [1]. Standard screens for phenotypes that relate to lignin defects, such as irregular xylem (*irx*) and reduced epidermal fluorescence (*ref*) have yielded initial mutants that are promising for biofuel applications [2]. However, more research is needed to obtain the necessary range and specificity of mutants for lignin engineering. To find novel genes involved in lignin biosynthesis, we are using suppressor and enhancer mutagenesis screens of known lignin mutants, an approach that so far, is unique and promises to yield genes not previously known to be involved in lignin biosynthesis. We are also using co-expression analysis on specific pathways and processes using publicly available, high-quality data sets from microarray expression experiments. Previously, co-expression experiments have been used successfully to explore novel genes involved in some aspects of cell wall biosynthesis [3]. However, co-expression analysis on the monolignol biosynthesis pathway has not been reported yet. There are many reasons to hope that these kinds of analyses may yield insights into as yet undescribed but fundamental aspects of lignification. For example, the cell biology of the lignification process is still very poorly understood. Recently, it has been found that two monolignol biosynthesis proteins, C3H and C4H interact at the endoplasmic reticulum membrane [4] but possible protein:protein interactions within and beyond the known monolignol biosynthesis enzymes has generally received little study. We are using protein complex purification to identify novel proteins that interact with lignin biosynthesis enzymes and that may therefore be involved in lignin biosynthesis

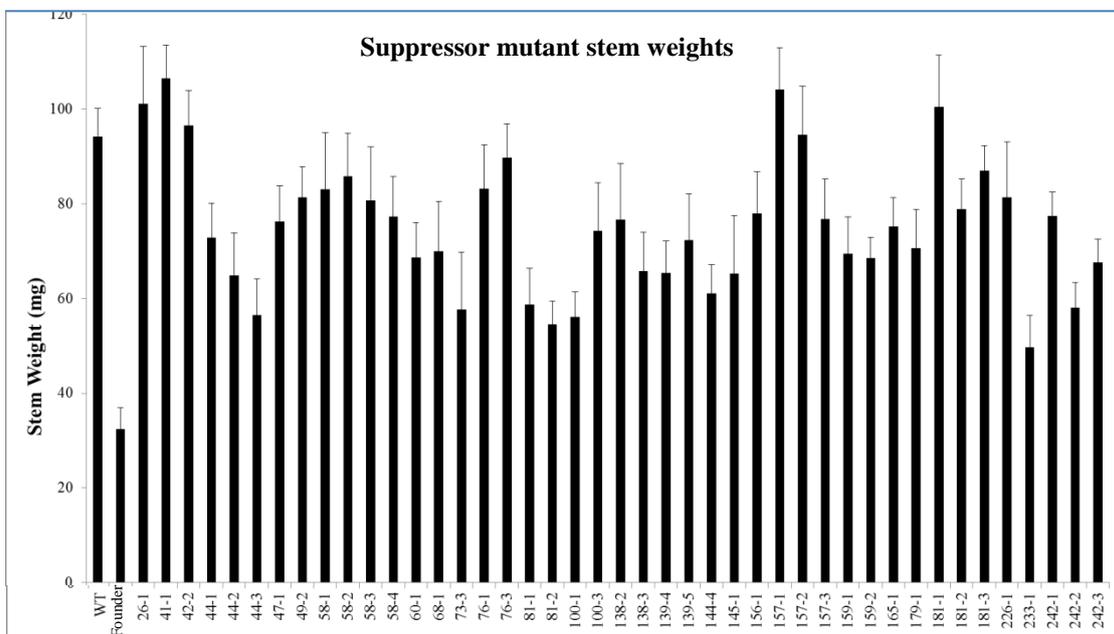
## **Results**

### *Mutant Screening*

To discover potential new target genes involved in improving saccharification properties in plants, a variety of approaches are being used in the model plant *Arabidopsis thaliana*. The first approach, a forward genetics approach, utilises genetic screening to identify mutants which show properties that are more desirable for the production of bioethanol. Three mutagenic screens have been performed. In each case, a founder mutant, which has a mutation in one of the monolignol

biosynthesis genes, was mutagenized with the aim of finding mutants that improve the original founder mutant's utility in the production of bioethanol. In two cases, the founder mutant shows a significant improvement of sugar release upon saccharification, but has reduced plant size. The aim of these two mutagenic screens therefore, was to find mutants which rescued the plant size defect, but still maintained the improved sugar release of the founder mutant.

The first screen, to find suppressors of the plant size defect, involved re-mutagenesis of a founder mutant with decreased lignin that releases more sugar upon saccharification of stem tissue, but has stunted growth [5]. The founder mutant was treated with EMS, a commonly used chemical mutagen. 80,000 plants from the resulting M<sub>2</sub> generation were screened for increased plant size. As reported last year, 71 mutants were identified originally, but upon re-screening, only 43 of these potential suppressors rescued the plant size defect of the founder mutant. For these 43 mutants, the stem weight of each mutant was quantified, since this tissue is the one most likely to be used in the production of second generation biofuels (Figure 1).



**Figure 1** – Stem weights of suppressor mutants isolated from a screen for increased plant size. *Arabidopsis* plants were grown for twelve weeks and stems collected. Each data point represents the mean of eight to ten stems, with error bars representing standard error of the mean.

While not all mutants were rescued back to the size of the wild type, every mutant identified has an increase in average stem weight that is statistically different from the founder mutant to a P value of 0.05 using a student's t-test, with the exception of two (mutants 73-3 and 233-1 in Figure 1).

While the mutants isolated showed an increase in plant size relative to the founder mutant, when analysis of the expression of founder gene in the potential mutants was performed, most mutants exhibited restored expression (Figure 2).

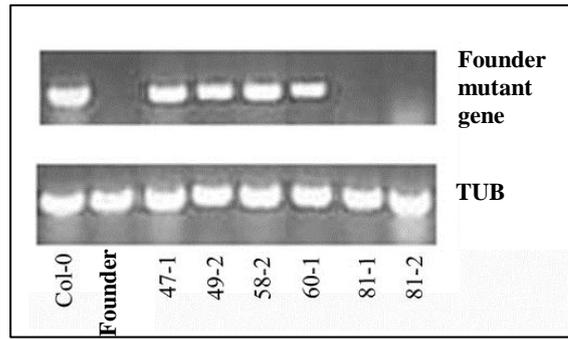
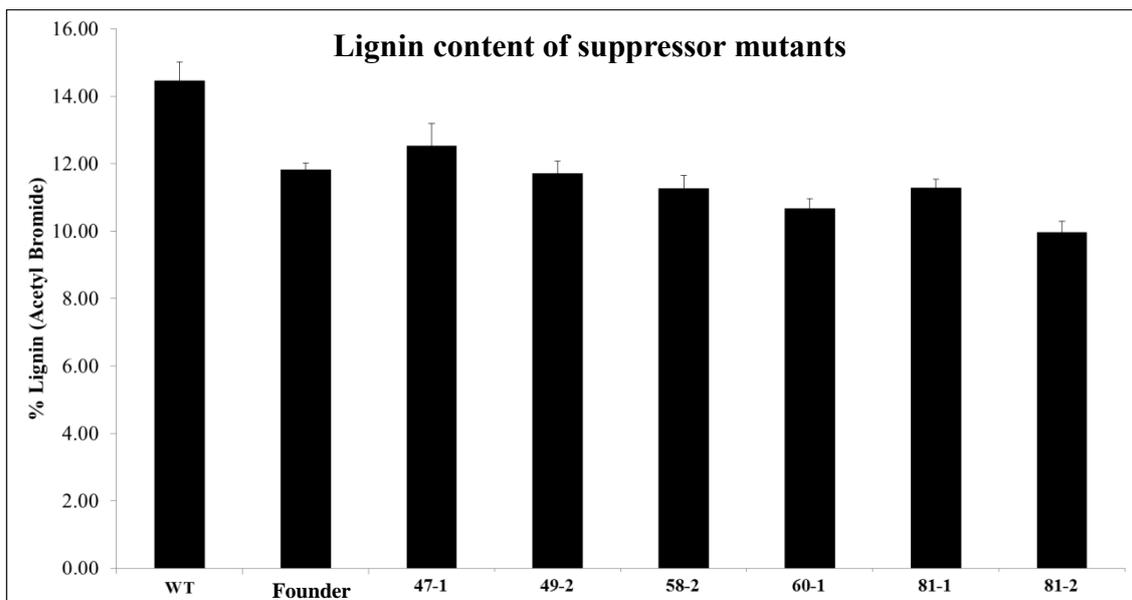


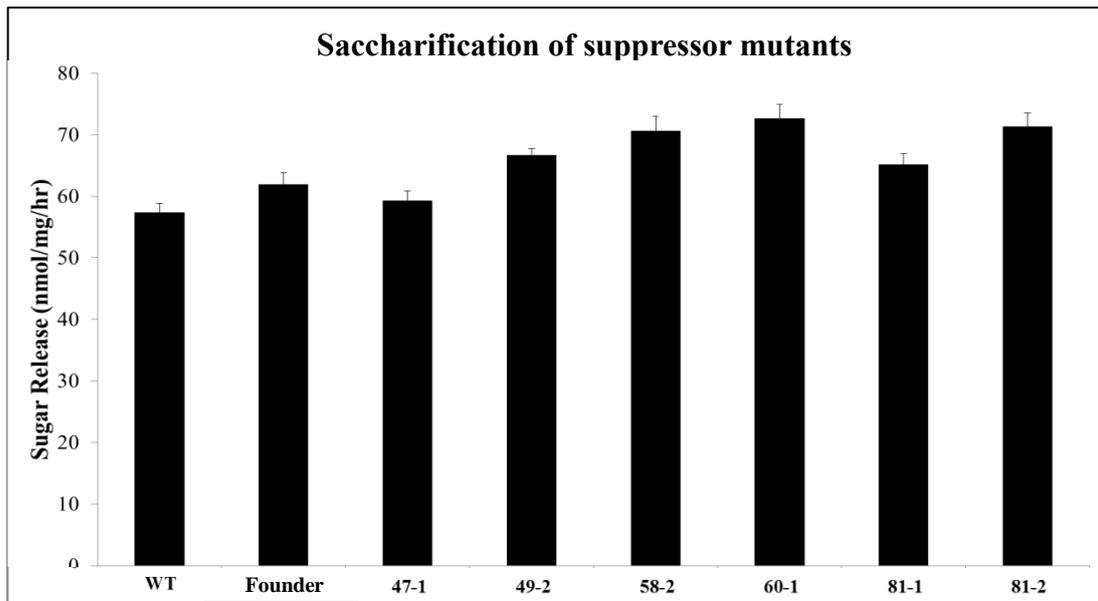
Figure 2 – RT-PCR of several suppressor mutants shows founder gene expression is restored in most mutants. RNA was extracted from stems of six-week old *Arabidopsis* plants.

Only three of the 43 mutants isolated (81-1, 81-2, 100-3) show no expression of the founder mutant gene when assayed by semi-quantitative RT-PCR, while five show reduced expression (44-1, 44-2, 44-3, 73-3, 100-1). Despite this restoration of expression, many of the mutant lines isolated still show reductions in lignin content when assayed by the acetyl bromide method (Figure 3) [6].



**Figure 3** – Acetyl bromide lignin content of several suppressors. *Arabidopsis* stems were collected after twelve weeks of growth and ground to powder. Powder was extracted four times with 80% ethanol and once with 2:1 chloroform: methanol. Extracted stem powder was treated with 25% acetyl bromide at 70°C for 30 minutes, and absorbance at 280 nm measured to quantify lignin [6].

All of the suppressor lines displayed in Figure 3 show reductions in lignin content which is significantly lower than wild type, with a P value of less than 0.05 in a student's t-test. Furthermore, when tested for saccharification, five of the six lines above showed significant differences from the wild type ( $P < 0.05$  in a student's t-test), with the exception being line 47-1 (Figure 4).



**Figure 4** – Saccharification analysis of suppressor mutants. Stems were collected from twelve-week old *Arabidopsis* plants and ground into powder. Powder was then used in high-throughput saccharification analysis as in Gomez *et al.* [7].

The reduction of lignin content and corresponding increase in saccharification yield suggests that, despite the restoration of expression of the founder mutant gene in many of the suppressors, there is still some defect in the gene product that prevents the plant from fully restoring lignin production.

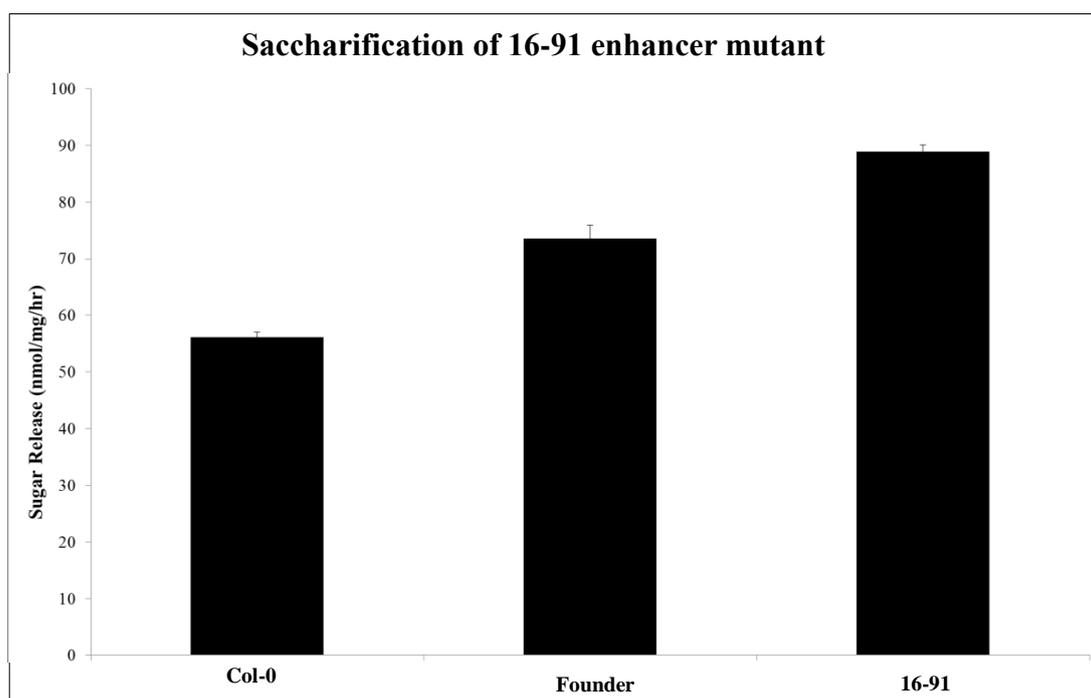
All six of the lines presented in figures 3 and 4 are now being used in mutation mapping. Each line has been crossed to the original founder mutant or to the wild type. Plants from the resulting F<sub>2</sub> generation that exhibit the rescued plant size defect have been selected for the mapping population. DNA from the mapping population will be isolated and a next generation sequencing approach will be used to identify the causal mutation, as developed by Abe *et al.* [8].

The second screen to find suppressors of the plant size defect was performed with the *ref3-1* mutant, which is mutated in the *C4H* gene of *Arabidopsis* [9]. This gene produces cinnamate 4-hydroxylase, an enzyme which participates in the production of monolignols. The *ref3-1* mutant has reduced lignin levels and increased sugar release upon saccharification, but has a significant plant size defect. Similar to the suppressor screen already described, the *ref3-1* mutant was mutagenized with EMS, and 80,000 M<sub>2</sub> plants were screened for an increased size phenotype. 25 potential mutants were isolated from the screen, but upon re-screening, none consistently exhibited the larger plant size phenotype. Therefore, this screen is not currently being pursued further as we already have too many promising candidates to identify from other screens.

A third mutagenic screen involves a different approach than used by the first two screens. In this screen, plants were screened for increases in sugar release upon saccharification. The founder mutant used for this screen is a lignin biosynthesis gene mutant which has small reductions in lignin and has no visible plant growth phenotype. The mutant also has an increase of sugar release upon saccharification. It was hypothesized that further increases in saccharification potential could be realised without affecting the growth of the plant. To find such plants, a high-throughput saccharification assay was used to screen a re-mutagenized ‘enhancer’ population [7]. This population was generated by mutagenizing the founder mutant with EMS, and

stem material from each member of the M<sub>2</sub> screening population was collected. 20,000 M<sub>2</sub> plants were grown for screening and 10,000 of these plants were used in a high-throughput saccharification screen. For each plant, stem material was collected and ground to powder for use in the saccharification procedure.

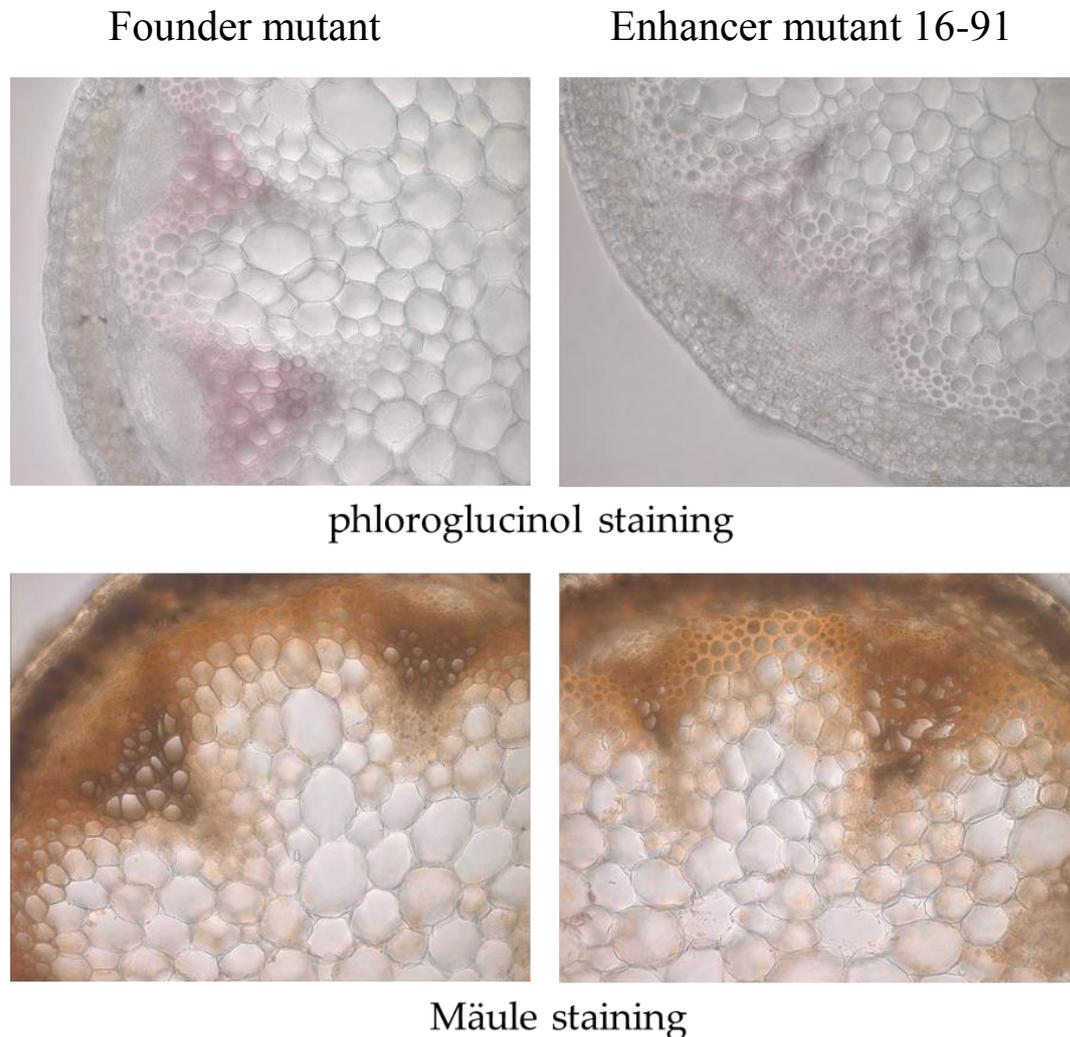
So far, approximately 4,750 stem samples have undergone the high-throughput saccharification assay, with the other 5,250 samples currently being assayed. Samples that were statistical outliers for sugar release in the saccharification assay were re-screened using the same assay. Out of 34 potential mutants that have been re-screened, one has shown a statistically significant increase in sugar release when compared to the founder mutant (Figure 5). Another 13 potential mutants currently await re-screening.



**Figure 5** – The enhancer mutant 16-91 shows an increase in sugar release upon saccharification. *Arabidopsis* plants were grown for twelve weeks, and stems were collected and ground to powder. Powder was used in the high-throughput saccharification assay developed by Gomez *et al.* [7]. Each data point represents four technical replicates of each of four biological replicates.

Characterisation of the 16-91 mutant is underway. Since increases in saccharification can be associated with decreases in lignin content, stem sections from the 16-91 mutant were stained with reagents to detect lignin (Figure 6). Interestingly, the 16-91 shows a marked decrease in phloroglucinol staining, an indication of decreased lignin content in the xylem of 16-91 mutant stems. The 16-91 enhancer mutant also displays a collapsed xylem phenotype (Figure 6), a phenotype indicative of cellulose, hemi-cellulose, or lignin deficiencies. Currently, the 16-91 mutant is being analysed for lignin content levels by the acetyl bromide assay. Furthermore, the 16-91 mutant has been crossed back to the wild type for mutation mapping. Similarly to the suppressor mutants from the first experiment, DNA from plants showing the enhanced saccharification phenotype will be isolated and used for the mapping

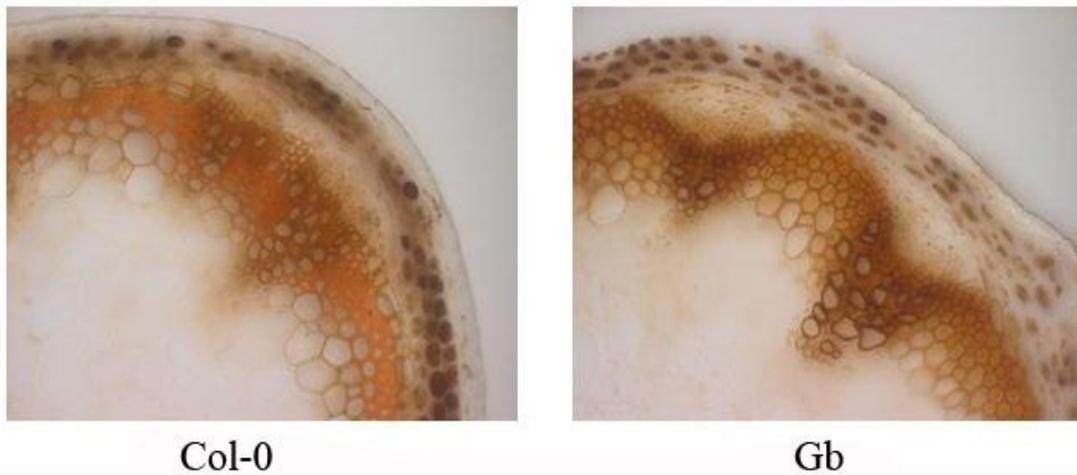
population. The MutMap program, which takes advantage of next generation DNA sequencing, will be used to find the causal mutation [8].



**Figure 6** – Lignin staining in stem sections of the enhancer mutant 16-91. Stems from five-week old *Arabidopsis* plants were collected and sliced into 100  $\mu$ m sections. Sections were stained with either phloroglucinol or Mäule reagent as in Corea *et al.* [1] and imaged with a light microscope.

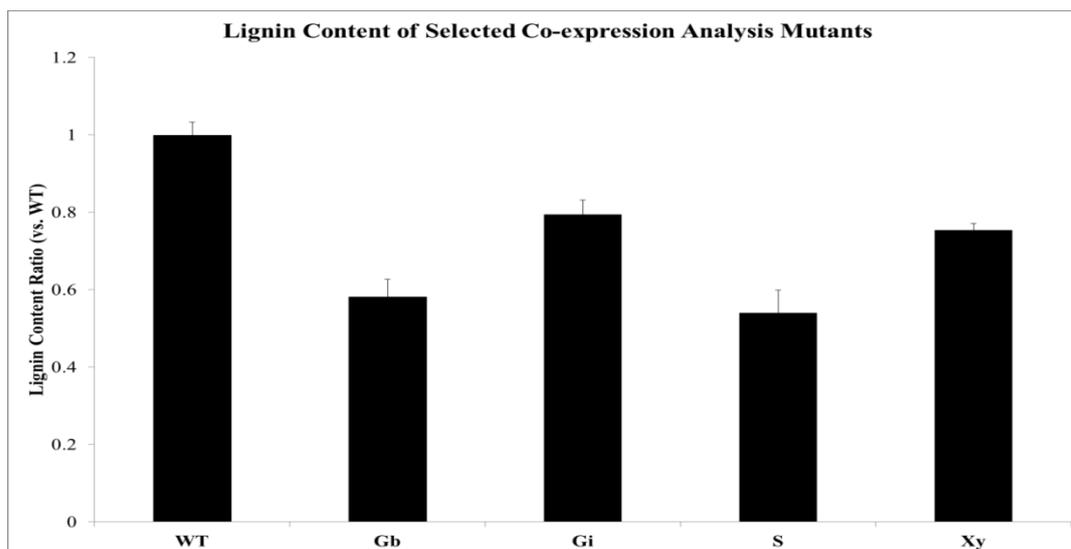
#### *Co-expression analysis and protein interactions*

Another approach used to identify novel genes involved in lignin production is gene co-expression analysis. In order to retrieve potential candidates involved in lignification, three individual co-expression analyses with known monolignol biosynthetic genes were performed: ACT [10], CressExpress [11] and Ray (Raymond Wightman and Simon Turner- personal communication). In total, 255 genes were retrieved, with some of them shared between different analyses. Through histochemical staining of lignin (phloroglucinol and Mäule staining), we were able to identify one mutant (Gb) exhibiting reduced S-lignin accumulation in inflorescence stem compared with wild type (Figure 7).

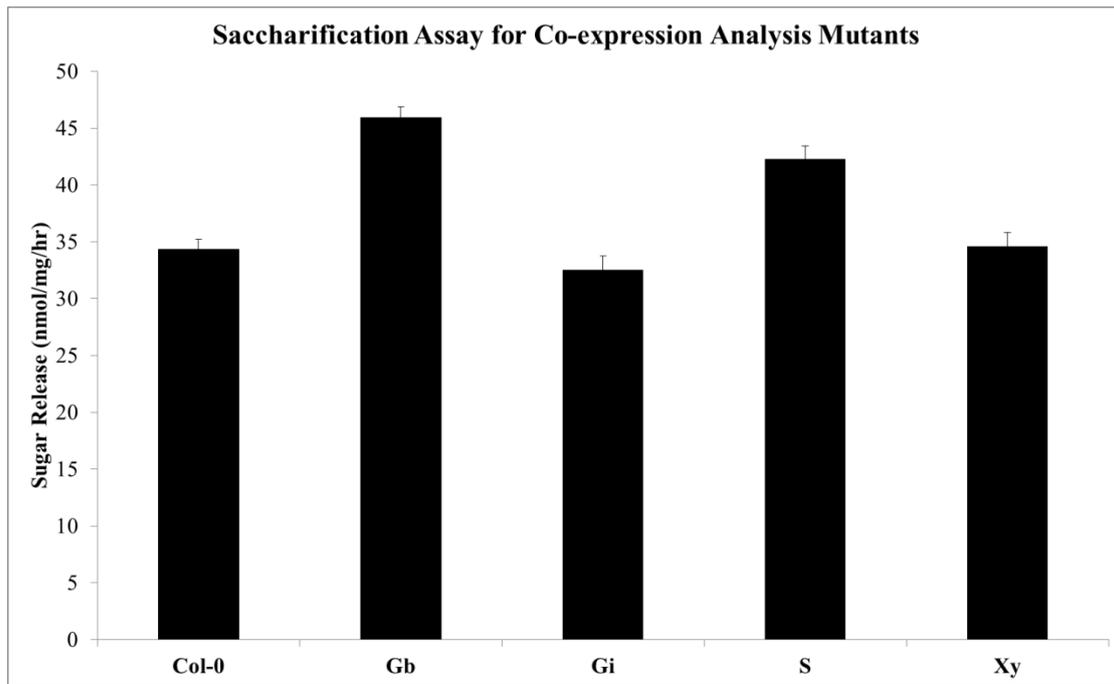


**Figure 7** – Mäule staining of wild type (Col-0) and Gb mutant stems. Plants were grown for six weeks, then stems were collected, sectioned, and stained with Mäule reagent.

We used acetyl bromide to more accurately quantify the lignin content of 66 lines mutated in genes identified from the co-expression analysis. Four mutants, including the Gb mutant, showed a decrease in lignin levels when compared to the wild type (Figure 8).

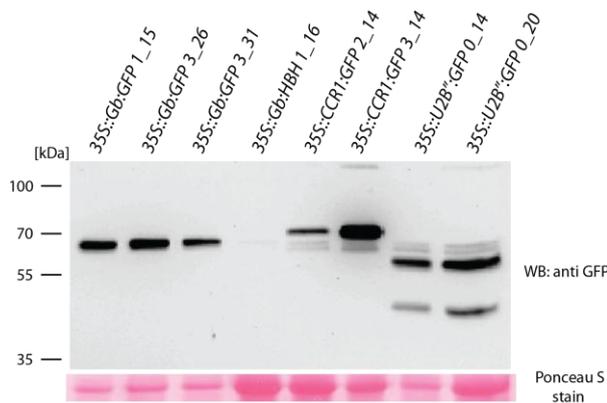


**Figure 8** – Acetyl bromide lignin content of co-expression analysis mutants. *Arabidopsis* stems were collected after 12 weeks growth and ground to powder. Powder was extracted four times with 80% ethanol and once with 2:1 chloroform:methanol. Extracted stem powder was treated with 25% acetyl bromide at 70°C, 30 minutes, and absorbance at 280 nm measured to quantify lignin [6].



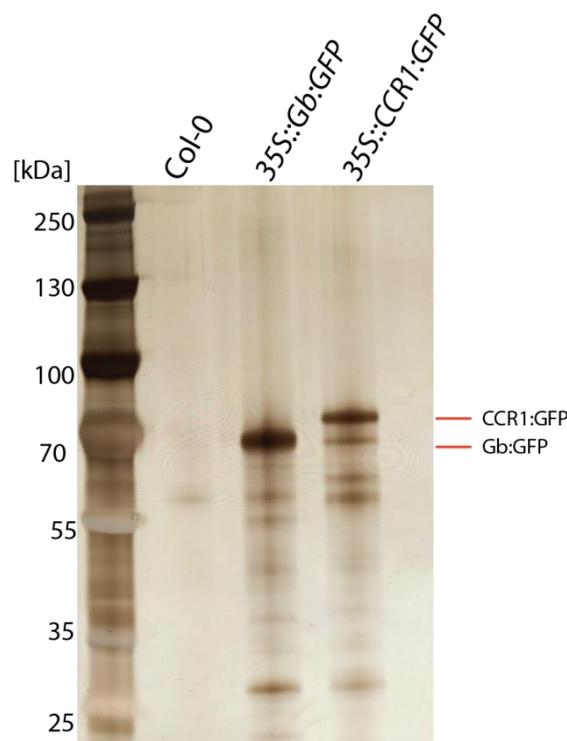
**Figure 9** – Saccharification of reduced-lignin co-expression analysis mutants. Stems were collected from 12 week old *Arabidopsis* plants and powdered for analysis in the high-throughput saccharification assay of Gomez *et al.* [7].

Analysis of protein complexes formed by known lignin biosynthesis enzymes could reveal new proteins involved in the lignification process. In order to identify protein interactions occurring *in vivo*, potential protein complexes in the lignin biosynthesis pathway are being purified directly from plants. *Arabidopsis* plants were produced that overexpress chosen lignin enzymes, and the Gb protein, fused to specific protein tags. Homozygous *Arabidopsis* transgenic lines were selected for all constructs. Presence of the tagged proteins in selected lines was confirmed using western blotting with tag specific antibodies (Figure 10).



**Figure 10** - Detection of GFP tagged proteins in homozygous *Arabidopsis* lines. Proteins were isolated from lines expressing Gb:GFP and CCR1:GFP. Plants expressing U2B''':GFP and Gb:HBH were used as a positive and negative controls. Proteins were separated on a SDS-PAGE gel and western blot was performed using a GFP antibody. Ponceau S staining was used to compare protein loading.

As shown in Figure 10, presence of GFP-tagged proteins in selected transgenic lines could be detected using a GFP-specific antibody. These transgenic lines were used in the next step for protein purification. In order to stabilise protein complexes present in the plants, a protein cross-linking step was first performed. This cross-linked material was used in protein purification. GFP-tagged proteins can be purified using a GFP trap, which consists of a small peptide that specifically binds GFP [12]. To purify complexes formed by the GFP-tagged proteins, a GFP trap fused with agarose beads was used. Presence of GFP-tagged proteins in the protein purification samples was confirmed by western blot. In order to visualize additional proteins co-purifying with proteins of interest, samples were analysed by silver staining (Figure 11).



**Figure 11** - Detection of proteins co-purifying with Gb:GFP and CCR1:GFP. Proteins were isolated with a GFP-trap from plants expressing Gb:GFP or CCR1:GFP fusions. Col-0 (WT) plants were used as a negative control. Proteins were separated on a SDS-PAGE gel and detected using silver staining.

Additional proteins co-purifying with Gb:GFP or CCR1:GFP proteins were detected using the silver staining method. These proteins could not be detected in the negative control, where no GFP tag was present. In order to identify all proteins co-purifying with selected tagged proteins, samples were analysed using liquid chromatography/mass spectrometry. Initial analysis of the proteins co-purifying with the Gb:GFP protein reveal the presence of known enzymes of the monolignol biosynthesis pathway. The presence of enzymes involved in lignin biosynthesis is consistent with a potential function of the Gb protein in this process. Further work is aimed at identifying exactly what role Gb plays in the lignification process.

### Progress

Several mutants from a variety of mutant screens have been identified in *Arabidopsis thaliana* that improve the saccharification properties of the plant. Some rescue a plant size defect of a desirable lignin-deficient mutant, while others increase

the saccharification potential of an existing mutant. These mutants have been quantified for lignin levels and saccharification, and mutation mapping is currently underway for several mutants in both screens. Gene co-expression analysis has identified two genes, that when mutated, produce mutants that contain less lignin and release more sugars upon saccharification. The protein produced from one of these genes has also been used for protein complex purification, along with previously studied lignin biosynthesis enzymes. Initial results indicate that the Gb protein can interact with downstream members of the lignin biosynthesis pathway, suggesting a novel role for this protein in monolignol biosynthesis and providing a new target for genetic manipulation in plants. A patent is being filed on this work in collaboration with Prof. Wout Boerjan, Gent University.

### **Future Plans**

For the remainder of the project, the focus for the mutant screening portion of the project will be on identifying the causal mutations behind the observed phenotypes. Currently, six lines from the first suppressor screen have been selected for mapping. In addition, the single mutant isolated so far in the enhancer screen is being prepared for mutation mapping. In addition, any mutants identified from the remaining samples to be processed in the initial screening will be characterised and prepared for mutation mapping. Work will continue to identify proteins interacting with lignin biosynthesis enzymes through LC/MS analysis of co-purifying proteins from prepared plant lines. A final major objective is to use insights from this and other work to try to illuminate the role of Gb in lignin biosynthesis.

### **Publications**

A publication on Gb is in preparation but will have more impact if we postpone submission until we have a clear idea of the role of Gb.

Papers on the suppressor and enhancer mutants must wait until we have identified the causal genes.

A patent application on 'Modified Plants' has been prepared and will be filed soon after submission to GCEP for approval.

### **References**

1. Corea, O.R.A., C. Ki, C.L. Cardenas, S.J. Kim, S.E. Brewer, A.M. Patten, L.B. Davin, and N.G. Lewis, Arogenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins, *J. Biol. Chem.*, 287, 11446-11459, 2012.
2. Li, X., N.D. Bonawitz., J.K. Weng, and C. Chapple, The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids, *Plant Cell*, 22, 1620-1632, 2010.
3. Brown, D.M., L.A. Zeef, J. Ellis, R. Goodacre, and S.R. Turner, Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics, *Plant Cell*, 17, 2281-2295, 2005.
4. Chen, H.C., Q. Li, C.M. Shuford, J. Liu, D.C. Muddiman, R.R. Sederoff, and V.L. Chiang, Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis, *Proc. Natl. Acad. Sci. USA*, 108, 21253-21258, 2011.
5. Mir Derikvand, M., J.B. Sierra, K. Ruel, B. Pollet, C.T. Do, J. Thévenin, D. Buffard, L. Jouanin, and C. Lapierre, Redirection of the phenylpropanoid pathway to feruloyl malate in *Arabidopsis* mutants deficient for cinnamoyl-CoA reductase 1, *Planta*, 227, 943-956, 2008.

6. Chang, X.F., R. Chandra, T. Berleth, and R.P. Beatson, Rapid, microscale, acetyl bromide-based method for high-throughput determination of lignin content in *Arabidopsis thaliana*, *J. Agric. Food Chem.*, 56, 6825-6834, 2008.
7. Gomez, L.D., C. Whitehead, A. Barakate, C. Halpin, and S.J. McQueen-Mason, Automated saccharification assay for determination of digestibility in plant materials, *Biotechnology for Biofuels*, 3, 23, 2010.
8. Abe, A., S. Kosugi, K. Yoshida, S. Natsume, H. Takagi, H. Kanzaki, H. Matsumura, K. Yoshida, C. Mitsuoka, M. Tamiru, H. Innan, L. Cano, S. Kamoun, and R. Terauchi, Genome sequencing reveals agronomically important loci in rice using MutMap, *Nat. Biotechnol.*, 30, 174-178, 2012.
9. Schilmiller, A.L., J. Stout, J.K. Weng, J. Humphreys, M.O. Ruegger, and C. Chapple, Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth, and development in *Arabidopsis*, *Plant J.*, 60, 771-782, 2009.
10. Manfield, I.W., C.H. Jen, J.W. Pinney, I. Michalopoulos, J.R. Bradford, P.M. Gilmartin, & D.R. Westhead, *Arabidopsis* co-expression tool (ACT): web server tools for microarray-based gene expression analysis, *Nucleic Acids Res.*, 34, W504-W509, 2006.
11. Srinivasainagendra, V., G.P. Page, T. Mehta, I. Coulibaly, and A.E. Loraine, CressExpress: A tool for large-scale mining of expression data from *Arabidopsis*. *Plant Physiol.*, 147, 1004-1016, 2008.
12. Rothbauer, U., K. Zolghader, S. Muyldermans, A. Schepers, M.C. Cardoso, and H. Leonhardt, A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins, *Mol. Cell. Proteomics*, 7, 282-289, 2008.

## Contacts

Claire Halpin: [c.halpin@dundee.ac.uk](mailto:c.halpin@dundee.ac.uk)  
Gordon Simpson: [g.g.simpson@dundee.ac.uk](mailto:g.g.simpson@dundee.ac.uk)  
Christopher McClellan: [christopher.mcclellan@hutton.ac.uk](mailto:christopher.mcclellan@hutton.ac.uk)  
Katarzyna Rataj: [katarzyna.rataj@hutton.ac.uk](mailto:katarzyna.rataj@hutton.ac.uk)  
Yuguo Xiao: [xiaoy@umn.edu](mailto:xiaoy@umn.edu)  
Abdellah Barakate: [abdellah.barakate@hutton.ac.uk](mailto:abdellah.barakate@hutton.ac.uk)  
Carinna Baird: [carinna.baird@hutton.ac.uk](mailto:carinna.baird@hutton.ac.uk)  
Sean Hackett: [sean.hackett@hutton.ac.uk](mailto:sean.hackett@hutton.ac.uk)  
Natalie Walker: [natalie.walker@hutton.ac.uk](mailto:natalie.walker@hutton.ac.uk)  
Lydia Welsh: [lydia.welsh@hutton.ac.uk](mailto:lydia.welsh@hutton.ac.uk)