Annual project report 2021

Project title	Exploiting variation in grain protein content and quality to determine effects of environment on processing quality		
Project number	21130058		
Start date	1 Oct 2019	**End date**	31 Jan 2023

Project aim and objectives

The studentship project will exploit genetic variation in wheat grain protein content and composition to explore the mechanisms for environmental effects of grain processing quality and to develop markers for increased stability. This builds on work at Rothamsted Research supported by AHDB and BBSRC.

1. It is known that Hereward is unusually efficient at converting fertiliser nitrogen into grain protein, a trait known as Grain Protein Deviation (GPD). Longer-term trials, as part of the Wheat Genetic Improvement Network (WGIN), show that Hereward also has higher stability of protein content compared with more modern cultivars. A population of 90 doubled haploid lines (RAGT) from the cross between Malacca (no GPD, low stability) x Hereward (high GPD, high stability) was used to carry out a genetic analysis of these two traits, using field trials in three environments (year x site).
   New high-density molecular maps of these lines will now allow these traits to be mapped in greater detail as quantitative trait loci (QTL). These regions will be aligned with new physical maps of the wheat genome revealing the gene content. Polymorphisms within these genes will be identified by exome capture, skim sequence, and transcriptome data. Putative causal polymorphisms will be converted to 'breeder friendly' molecular marker such as KASP and validated.

2. A previous study of the Malacca x Hereward lines identified QTLs that affect breadmaking quality (e.g. loaf volume and crumb structure). The 'good alleles' at these 'quality' QTLs are derived from both parents and sets of near isogenic lines (NILs) will be used to explore their mechanisms without the confounding effects of differences in genetic background.
   The NILs will be used to determine the impacts of the environment on individual aspects of protein quality, identifying differences in stability of quality between alleles and QTLs, determining mechanisms and establishing markers for stable high-quality alleles. This will be achieved by growing sets of lines with the good and poor-quality alleles at 6 QTLs in replicate field trials in four environments (JIC and Rothamsted over two years). The sets of lines will be bulked for analysis of their processing quality - Milling, Farinograph, Extensograph and test baking analyses. Associated differences in protein composition will be determined by characterising glutenin polymers (major determinants of dough strength) by SE-HPLC.

Annual project report 2021

Key messages emerging from the project

1. Comparison of the sets of NILs for quality QTLs has confirmed that the 4D-2QTL has a reproducible impact on the texture of bread made from white flour. This is, therefore, a potential source of improved quality for breeders and will be studied in detail to determine the mechanism

2. Grain protein deviation (GPD) is a reproducible source of increased grain protein without requiring additional nitrogen fertilisation. The identification of QTLs and molecular markers will, therefore, provide an important source for greater sustainability of production of breadmaking wheat.

Summary of results from the reporting year

1. Comparison of NILs for QTLs affected baking quality

Six sets of NILs with good and poor QTLs associated with breadmaking quality traits were grown in 2019-2020 at Rothamsted (Table 1). Each set is formed of five lines with the Hereward allele at the QTL locus and 5 lines with the Malacca allele at the QTL locus, making a total of 10 sets of lines.

Recurrent Parent	QTL chromosome	QTL quality Trait
Malacca	1B	Cell number
Malacca	2D	Loaf volume
Malacca	4D-1	Cell number and Crumb whiteness
Hereward	4D-2	Crumb whiteness
Malacca	6A	Cell number
Hereward	7B	Loaf volume

Table1. Description of the six NIL sets and the QTL location and trait they are associated to. The Recurrent parent is the parent used for the backcross.

1.1. SE-HPLC

A good quality bread requires dough that exhibits a balance between elasticity or dough strength for gas retention and extensibility to allow the expansion and to achieve a good loaf volume.

The gluten protein and prominently the glutenins present in the white flour confer these essential properties. Glutenins are large polymeric proteins composed of a mixture of high and low molecular weight subunits. SE-HPLC separates the gluten protein into 5 peaks: the high molecular weight

glutenin polymers or F1 fraction, the low molecular weight glutenin polymers (F2), the high molecular weight gliadins (F3), the low molecular weight gliadins F4, and the albumin/globulin fraction (F5).

The ratio between HMW-GP and LMW-GP (F1/F2) and the HMW-GP fraction (F1) correlates positively with dough strength, and that the ratio between gliadin and glutenin (F3+F4)/(F1+F2), and the ratio gliadins/HMW-GP (F3+F4)/F1) correlates positively with dough extensibility.

Analysis of SE-HPLC fractions

For the QTLs 1B, 4D-2 and 6A the favourable alleles (1B-b, 4D-2-a, and 6A-b) are associated with a higher F1/F2 ratio, whereas for QTL 2D and 4D-1, the favourable alleles (2D-b and 4D-1-b) are associated with a lower ratio F1/F2, while no difference in ratio was observed between the QTL7B alleles. As the ratio F1/F2 correlates positively with dough strength, bulks showing differences in this ratio may show differences in dough strength.

The ratios (F3+F4)/F1 and (F3+F4)/(F1+F2) correlate with extensibility. For the QTL 2D, 4D-1, and 7B, the favourable alleles (2D-b, 4D-1-b, and 7B-a) were associated with a higher ratio (F3+F4)/F1, whereas, for the QTL 1B, 4D-2 and 6A the favourable alleles (1B-b, 4D-2-a, and 6A-b) were associated with a lower ratio (F3+F4)/F1.

Correlation with Quality Parameters

Breadmaking

Analyses of the breadmaking quality of flours from the sets of lines, carried out by Heygates Ltd., confirmed that the 4D-2 QTL has a reproducible and statistically significant effects on the quality of bread.

Rheology

Differences in dough properties were determined with a Farinograph (dough stability and peak time) and an Extensograph (resistance to extension). This showed that the differences in (F3+F4)/F1 and (F3+F4)/(F1+F2) were related to extensibility determined using an Extensograph. The dough strength ranged from 205 to 248 BU for resistance to extension, and from 3.3 to 3.4 and 3.8 to 4.1 min for dough stability and peak time, respectively. The differences were smallest (4 BU) for QTL 2D and greatest for QTL 1B (27 BU) for resistance to extension, smallest for 2D (no difference) and greatest for 4D-1 (0.7 min), and smallest for 1B (no difference) and greatest for 2D and 4D-1 (6 s). Interestingly, the different measurements do not agree with each other. For instance, QTL1B has the highest difference for resistance to extension and the lowest difference for peak time. Figure 1 shows correlations between SE-HPLC fractions and Extensograph and Farinograph parameters. F1 shows a strong positive correlation with the ratio F1/F2 and negative correlations with both ratios (F3+F4)/F1 and (F3+F4)/(F1+F2). Surprisingly, F1 also appears to be strongly negatively correlated with F3 (r=-0.77). This shows that the proportions of HMW glutenin polymers and HMW gliadins (which are mainly omega-gliadins) are negatively related.

The ratio F1/F2 and the fraction F1 (HMW-GP) are not significantly correlated with stability (stab) or peak time (p), both measures of dough strength. However, the parameters are negatively correlated with dough resistance (a measure of dough strength) determined by Extensograph.

The ratios (F3+F4)/F1, (F3+F4)/(F1+F2), and the Fraction F4 (LMW-gliadins) all correlate negatively and significantly with the Extensograph extensibility.

1.2. Metabolomics

It has been reported that the addition of 1% and 2% wheat arabinogalactan peptide (AGP) to flour resulted in an increase of dough maximum resistance to extension, a decrease of dough extensibility and a slight decrease in loaf volume.

We determined the contents of AGP in the sets of NILs using ^{1}H-NMR spectroscopy. This allows AGP to be determined as arabinose and galactose equivalents. The results showed variation between the QTLs and between the alleles at each QTL (Figure 2). The values range from 3.2 nmoles/mg for line 186-8-6 at QTL 7A-b to 5.9nmoles/mg for line 207-6-22 at QTL 4D-2a with an average of 4.2 nmoles/mg. The content of AGP measured as galactose equivalents shows as a similar range, from 3 nmoles/mg for line 186-8-6 (QTL 4b) to 5.6nmoles/ng for line 207-6-22 (QTL 4D-2a).

One-way Analysis of Variance (ANOVA) was performed to determine differences between the allelic pairs. This revealed a highly significant QTL effect ($p<0.001$) for all measurements, indicating that at least 2 QTL means differ significantly from each other. However, considering comparison of alleles within the QTL, no significant differences at $p=0.05$ were observed for the three traits, although the allelic difference in arabinose content at QTL2D was almost significant ($p=0.09$).

Annual project report 2021

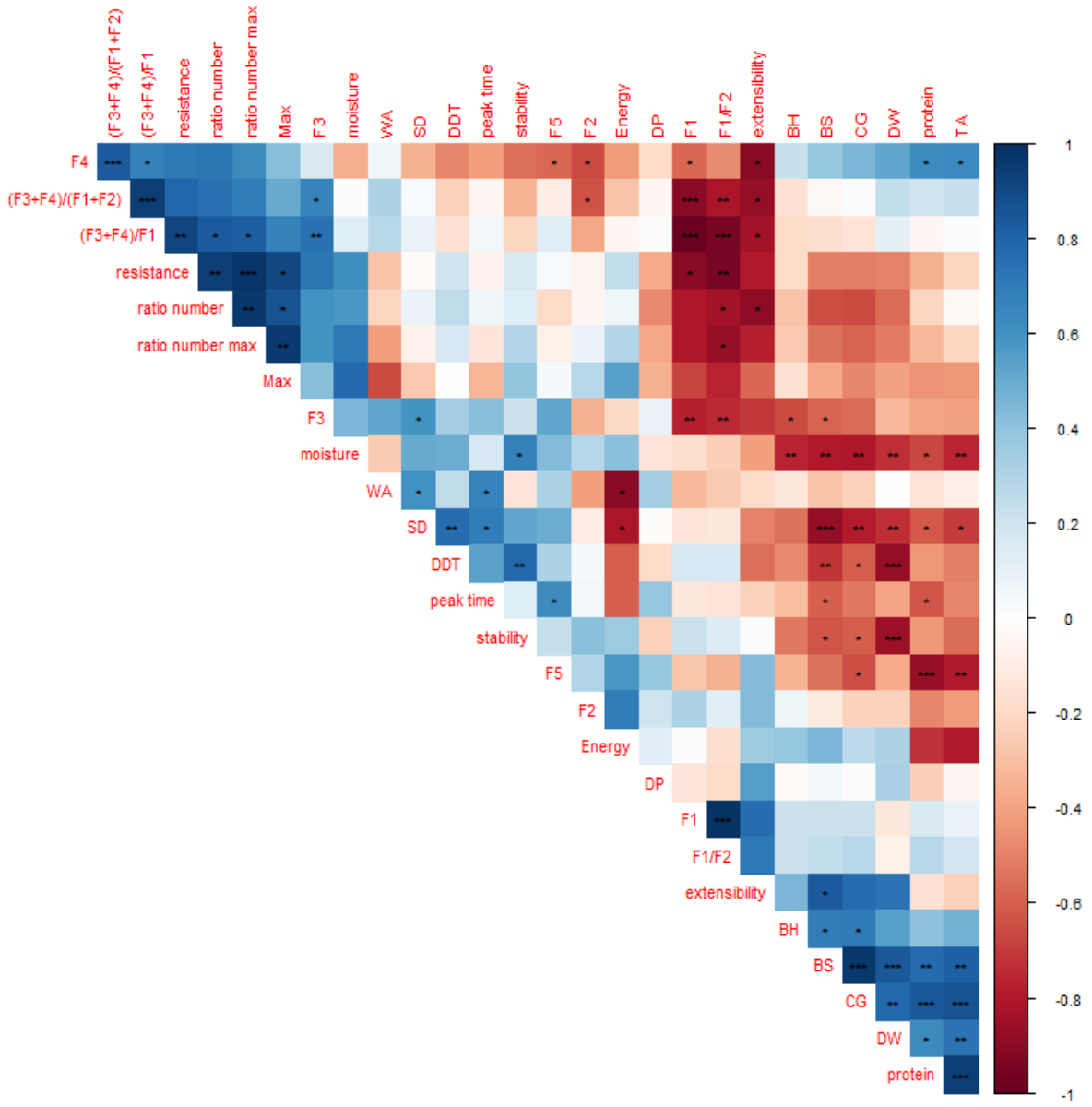

Figure 1. Correlation matrix including SE-HPLC ratio, rheology and baking data. BH: bake height, DW: Degree of Weakening, CG: colour grade, BS: Bran Scan, WA: water absorption, SD: starch damage, DDT: dough development time, DP: dough proof.

The colour of the tiles indicates the relationship between the variables, blue for a positive and red for a negative relationship. The colour intensity relates to the strength of the correlation; dark red correlations are higher than light red. Significant correlations ($p<0.05$) are marked with one ($p<0.05$), two ($p<0.01$) or three stars ($p<0.001$).

The results described in this summary report are interim and relate to one year. In all cases, the reports refer to projects that extend over a number of years.

Annual project report 2021

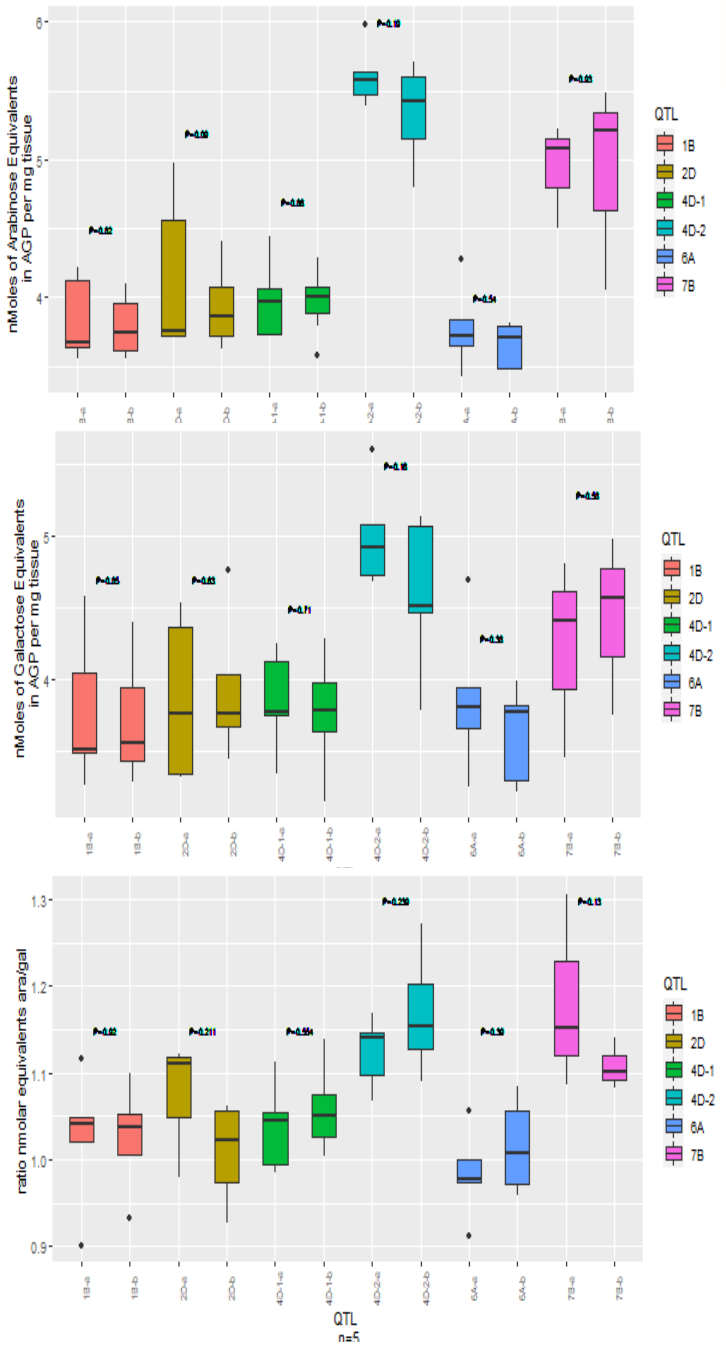

Figure 2. [1]H-NMR spectroscopy to quantify a: Arabinose equivalents, b: Galactose equivalents, and c: ratio Arabinose:Galactose equivalents in AGP per mg of tissue. n=5. The p-values are the result of a contrasted ANOVA model. Six contrasts were designed for pairwise comparison of the allelic pairs.

2. Grain protein deviation (field trial 2019–20)

2.1. Protein content and yield

Figure 3a summarises the protein content (% dry wt.) of the 300 double haploid lines. The median protein content is of 12.81%. Similarly, the yields of the 300 lines were highly variable ranging from 593.6 $g.m^{-2}$ to 1397.8 $g.m^{-2}$ with an average of 1031.9 $g.m^{-2}$ (Figure 3). To calculate GPD values, a simple linear regression was fitted between grain protein content and grain yield (Figure 4). There is a wide variation in both protein content and yield. The regression slope is negative and its coefficient close to zero, but statistically significant (β=-0.0027, $p<0.001$). The R^2 is weak (0.08), explaining only 8% of variation in both protein content and yield.

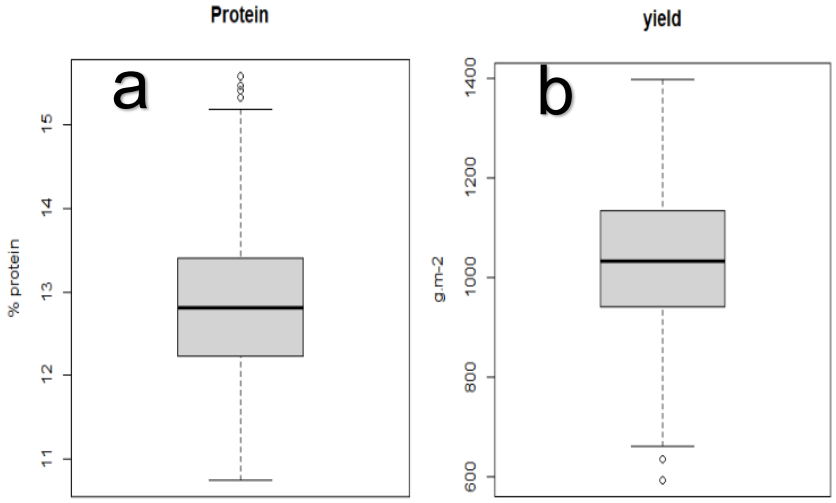

Figure 3: Boxplots of protein content (a) and yield (b) for the 300 double haploid lines. The solid dark lines mark the median, the boxes correspond to the interquartile range (Q3-Q1) and the dashed lines are the whiskers. The circles highlight potential outliers.

2.2. Calculation of GPD

A mixed linear model was used to account for the field structure and correct imbalance of the treatment effects. The GPD values vary from -0.90 to 0.89 with an average of -0.03. Regarding the size of the GPD, Shewry *et al.* (2013) found GPD values of 0.05 and 0.10 for Hereward and of -0.05 and +0.05 for Malacca for two field trials grown at Rothamsted in 2009 and 2010. In our case, Hereward has a GPD of 0.64 while Malacca has a GPD of -0.75. Thus, our GPD values seems stronger that the ones reported by Shewry *et al.* (2013). It is not surprising as the GPD is strongly affected by the environment and, therefore, varies across field trials.

Annual project report 2021

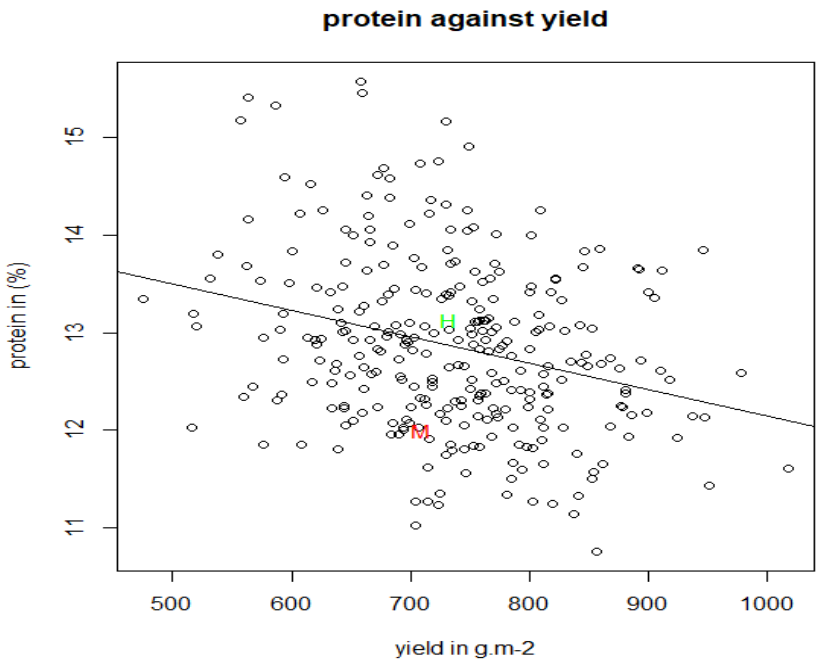

Figure 4. simple linear regression between grain yield (x) and protein content (y). The averages for the parents Hereward and Malacca are highlighted in green and red, respectively.

Key issues to be addressed in the next year

The final year of the project will focus on the following:

1. Confirm the genetic analysis of GPD by analysing data from the third field trial and combining the dataset with those for the first two experiment. This will allow the identification of molecular markers for breeders and the future identification of candidate genes.

2. Combine the datasets from analysis of the DH population and combine these with data from the Breeding and End Use LINK project to identify grain metabolites associated with grain quality QTLs.

3. Complete analyses of the sets of NILs, focusing on the 4D-2 QTL which affects bread texture.

Lead partner	Peter Shewry, Rothamsted Research
Scientific partners	John Innes Centre University of Reading
Industry partners	Heygates Ltd.
Government sponsor	

Annual project report 2021

Has your project featured in any of the following in the last year?	
Events	***Press articles***
Conference presentations, papers or posters	***Scientific papers***
- Poster presentation at the conference Monogram 2021 - Talk at Monogram 2022 conference - Presentations at PhD symposia at Reading University and Rothamsted Research	
Other	