Project Report No. 429

March 2008

Price: £8.50

Investigating wheat functionality through breeding and end use (FQS 23)

by

S J Millar[1], J Snape[2], J Ward[3], P Shewry[3], P Belton[4], K Boniface[5] and R Summers[6]

[1]Campden and Chorleywood Food Research Association, Station Road, Chipping Campden, Gloucestershire GL55 6LD
[2]John Innes Centre, Norwich Research Park, Colney Lane, Colney, Norwich, Norfolk NR4 7UH
[3]Rothamsted Research, West Common, Harpenden, Hertfordshire AL5 2JQ
[4]University of East Anglia, Norwich, Norfolk NR4 7TJ
[5]ADM Milling Ltd, Technical Centre, Hosegoods Mill, Old Dock, Avonmouth, Bristol, Avon BS11 9DL
[6]RAGT Seeds Ltd, The Maris Centre, 45 Hauxton Road, Cambridge, Cambridgeshire CB2 2LQ

This is the final report of a Defra-sponsored Link Project (FQS 23) that ran for six and a half years from August 2001. The project was sponsored by Defra (£285,309), BBSRC (£269,527) and HGCA (£215,219, Project No. 2492) with industrial support from nabim (£40,000) and the Federation of Bakers (£5,000). There were also in-kind contributions from industry (£94,276 from RAGT Seeds Ltd, £87,085 from Nickersons, £77,314 from Syngenta Seeds and £28,000 from the National Association of British and Irish Millers) making a total of £1,101,730.

Table of Contents

ABSTRACT

The UK wheat supply chain has seen significant changes in the last 40 years with a move from a market dependent on imports to one where the majority of wheat for a range of uses is home-grown. With increasing requirements and competitiveness in the marketplace, ongoing improvements in wheat yield, robustness and quality are fundamental to further improvements in the supply chain.

This project was initiated with the aim of generating new links between wheat genetics and processing performance to give breeders clear targets for development of high value varieties and to give the UK grain chain a significant step forward in overall competitiveness.

Three new wheat recombinant populations were developed and a wide range of genetic markers applied to each to develop a genetic map. Grain was produced from these populations over two growing seasons (2005 and 2006) and fully characterised in terms of milling and baking performance. Four bakery products were produced which represent the main commercial outlets for high quality wheat as well as offering insights to the technical aspects of interactions during processing: Chorleywood Bread Process (CBP) white bread, CBP wholemeal bread, no time dough (Spiral) white bread and puff pastry. New methods to objectively analyse the characteristics of these products were developed and validated by the project consortium allowing ranking of samples to be achieved using common and transferable techniques.

A wide range of new quantitative trait loci (QTL - genetic regions or loci where variation is correlated with an aspect of grain characteristics) were identified for specific aspects of processing quality such as milling yield, loaf volume, crumb colour and softness. These will now be used by plant breeders to generate varieties with better and more consistent processing quality than those currently available.

This project has represented a unique opportunity for representatives from across the supply chain to work together to generate improvements for all. By using end product quality as the target response, the project has provided wheat breeders with clearer targets for specific end uses, will lead to growers and processors having wheat better-adapted for their needs and the cereal science community with a major advance in the understanding of the genetic control of wheat quality.

PROJECT SUMMARY

Introduction

The last 40 years have seen dramatic changes in the wheat supply chain in the UK with a significant increase in the usage of home-grown wheat. Currently, some 85% of the 5.5M tonnes of wheat used by the UK flour milling is produced domestically, which represents double the proportion found in the 1970s.

A significant proportion of UK milling wheat is traded on a specified varietal basis which underpins the ability of the supply chain to produce flour having the correct characteristics for its intended end use as well as to ensure that appropriate financial value is assigned. The development of new varieties is key to the entire supply chain as well as being critical to the success of individual breeding companies. The system for adoption of new varieties is, however, highly competitive and necessitates large breeding programmes to ensure that breeders afford themselves the best possible chance of generating commercially viable advances in varietal characteristics. The time scale inherent in such strategies means that wheat breeding contains significant risk of failure, often after many years of work. As a result, there is a strong emphasis in generating more rapid ways of bringing varieties to market as well as ensuring that key characteristics may be efficiently adopted within breeding programmes.

Over the last years, breeders have been successful in developing markers for a number of characteristics affected by major genes which are used for selection on both agronomic and quality grounds. There is a recognition, however, that many of the characteristics related to milling and baking performance are as yet not characterised and in many cases are likely to be the result of a number of minor genes.

These considerations led to a series of discussions in 2000 and 2001 which culminated in the development of this project. It was recognised that advances in breeding at that time would facilitate the more rapid development of breeding populations and that this meant that a single project could include development of appropriate lines as well as their assessment. In addition, work on objective methods for assessing baked product quality was commencing which offered a means of giving breeders more robust methods of characterising overall performance for lines in a given population.

The project was conceived, therefore, as a means of developing new recombinant populations to which a range of genetic markers could be applied and against which a full range of objectively assessed milling and baking parameters could be used to determine new quantitative trait loci (QTL) for key aspects of wheat processing quality. As a means of underpinning this work, subsets of the populations developed were also assessed by newly-developed methods (FT-IR of hydrated wheat flour and metabolomics) to assess aspects of the underlying grain biochemistry.

The objectives addressed by the project were:

- To improve the understanding of raw material functionality and processing (primarily in relation to end use qualities of milling and breadmaking) by creating a benchmark "quality map" using current breadmaking wheat varieties and a combination of existing data plus novel quality parameters.

- To facilitate matching end users' requirements to appropriate raw material through targeted wheat breeding via the development of genetic maps of genes for processing quality.

- To evaluate experimental populations of doubled haploid lines produced from established parent varieties representing "key" variation in UK winter wheat germplasm for processing quality traits.

Materials and methods

Three recombinant doubled haploid populations were developed using hard milling, breadmaking varieties representing a range of quality characteristics: Malacca × Charger, Hereward × Malacca and Shango × Shamrock. These populations were used to produce ~25kg quantities of wheat from each of the harvest years 2005 and 2006.

Genetic maps for each of the populations were developed using Simple Sequence Repeats (SSR) and Diversity Array Technology (DArT) markers.

Each of the bulk samples was milled to produce white and wholemeal flour. These flours were then used to produce Chorleywood Bread Process (CBP) white bread, CBP wholemeal bread, no time dough (Spiral) white bread and puff pastry. A range of

wheat, flour and baked product characteristics were assessed using objective methods, many of which were developed specifically within the project. The data thus generated were then used for discovery of new QTL for each population. New QTL were identified using a range of approaches including single marker ANOVA and marker regression approaches. New QTL were evaluated in terms of significance as well as robustness across the 2 harvest years to enable selection of those likely to have most impact.

For each product and population, samples were ranked according to their performance as determined using a selection of agreed measures. This ranking was used to identify those samples representing the upper, middle and lower performing samples for a given population × product combination. These samples were then used as the basis for the underpinning work by FT-IR and metabolomics. In the former case, samples were hydrated using either H_2O or D_2O and spectra were collected using a single reflection ATR (attenuated total reflectance) cell. Metabolite data were assessed using ^1H NMR, the spectra from which were scaled and peak regions segregated prior to principal components analysis (PCA) and correlation analysis.

Results and discussion

Initial studies using data generated through annual new variety testing demonstrated that while routinely-measured quality parameters gave a useful picture of flour suitability for particular end uses, their ability to reliably predict functionality was limited. This served to underline the importance of the approach undertaken within the project whereby links between objective measures of milling/baking performance and underlying wheat genetics were targeted.

The objective methods of samples assessment were evaluated by using ring trials of real bread samples as well as through the use of newly-developed standards to facilitate loaf volume assessment. These exercises ensured that individual laboratories were able to make any necessary changes to methodology to ensure consistency of results. In addition, the project also included extensive use of C-Cell, a new image analysis-based system for the assessment of the structure of aerated products. This was seen as a major advance over more subjective scoring methods which had been used previously, both in terms of comparison between products as well as consistency over time.

An example of the range of product characteristics is given in Table 1 which represents a summary of basic grain quality assessments for the 3 populations over both years.

Table 1. Basic wheat quality parameters from the sample set used in the project for QTL analysis

	Protein content (%) dmb		**Grain hardness index by NIR**		**Hagberg Falling Number (s)**	
	Mean	Range	Mean	Range	Mean	Range
	2005 harvest					
H×M	14.4	13.5-15.2	72	56-83	322	169-418
M×C	13.1	11.8-14.8	60	45-70	351	236-440
S×S	12.8	12.1-13.8	70	44-93	211	98-302
	2006 harvest					
H×M	14.6	13.5-15.7	65	57-74	405	352-443
M×C	13.4	12.2-14.4	65	47-81	417	288-492
S×S	15.1	13.9-16.9	77	61-93	385	312-450

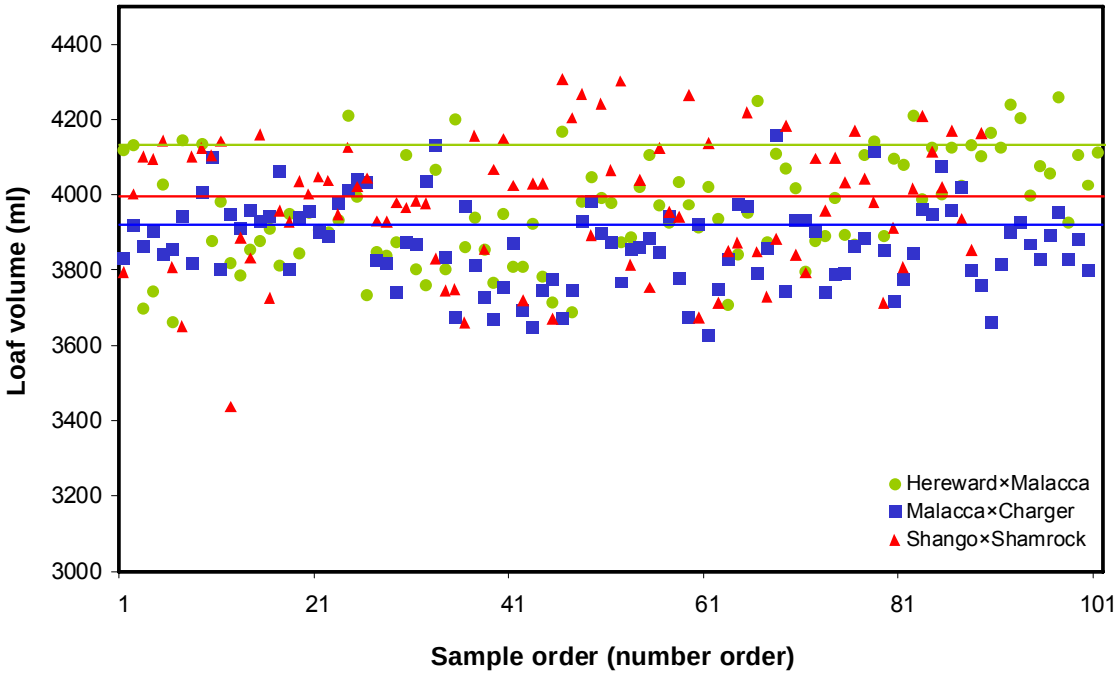
H×M = Hereward × Malacca population, M×C = Malacca × Charger population, S×S = Shango × Shamrock population

It can be seen that each of the populations showed reasonable ranges of response for each of the parameters, indicating that there was a good chance that new QTL could be identified. However, it was also clear that there were significant differences in grain quality between the 2 years with the results for 2005 harvest showing a tendency to lower Hagberg Falling Numbers. This was primarily due to the dry period experienced during early harvest in 2006. The samples from 2006 also showed increased protein content for the Shango × Shamrock population when compared with those from 2005 and this also had an effect on overall baking performance.

When baking responses within populations were assessed, the range of performance for individual lines was clearly shown. Figure 1 shows loaf volumes for bread produced using no time dough (Spiral) procedures from 2006 harvest wheat. The lines across the figure represent the upper value for the parents in each of the

populations. It can be seen that there are a number of lines in each population which exceed this upper parental value. This is an example of transgressive segregation where individual lines represent values outside the range of the parental values. This is very important in demonstrating that there is significant potential for the development of new varieties which out-perform currently-available wheat for key quality criteria.

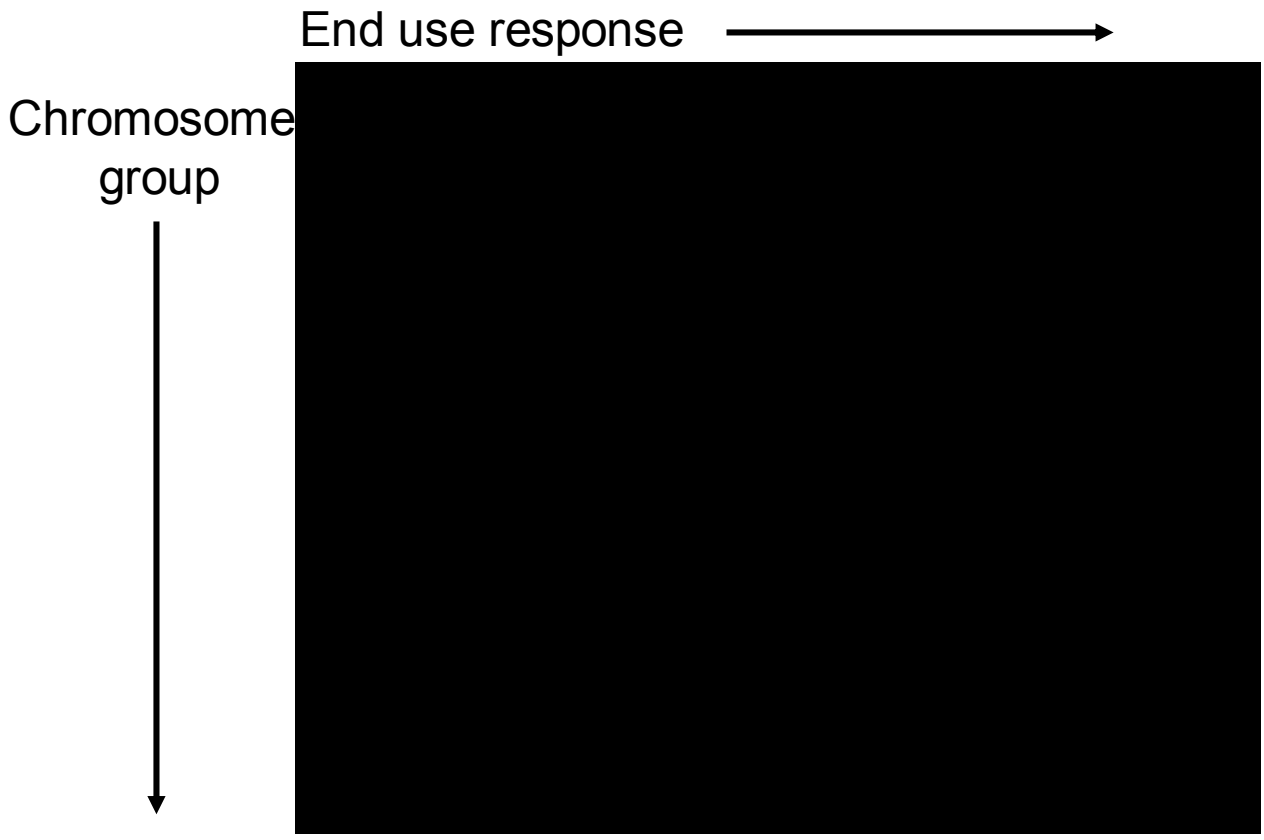
Figure 1. Loaf volume for Spiral white loaves for all populations from the 2006 harvest



A total of 606 new QTL for a wide range of breadmaking characteristics were identified across the 3 populations, representing an important new resource for plant breeders in developing improvements within UK wheat breeding programmes. Some of these QTL may represent the pleiotropic effects of the same underlying genes. In determining which QTL have the most potential, comparisons between years have been an effective means of highlighting those which are the most robust. A series of tables having the form of Figure 2 have been generated for each population. The specific results for each of these are confidential to the project consortium but Figure 2 represents an anonymised example where coloured blocks represent different levels of statistical significance for new QTL. The chromosomes involved are annotated going down the table while the different parameters assessed are given from left to right.

The number and range of QTL determined as a result of this represent a major advance for UK plant breeders and give an opportunity to look at 'stacking' positive QTL in new varieties such that significant improvements in processing quality may be envisaged.

Figure 2. Example format used in reporting QTL data to members of the project consortium



As a first stage of the metabolomics analysis, principal components analysis (PCA) was used to group the NMR data and to determine the spectral differences between the various groups. Differences were observed for the different varieties as well as some indication of environmental effects based on different grouping for the Malacca samples from 2 different populations. In addition, the NMR data were also used to develop correlations with data generated through milling and baking trials. A multivariate statistical technique, orthogonal partial least squares (OPLS), was used to develop models against these parameters. The results for CBP loaf volume, in particular, were interesting with statistically significant correlations being found between chemical shifts in the NMR data and the baking response (Figure 3). Given the historical and ongoing research investigating the impact of polymer structure on

dough properties, it was interesting to see evidence of variation in metabolite profile (relatively small molecules) on functionality.

Figure 3. Examples of chemical shifts showing positive and negative correlations to loaf volume

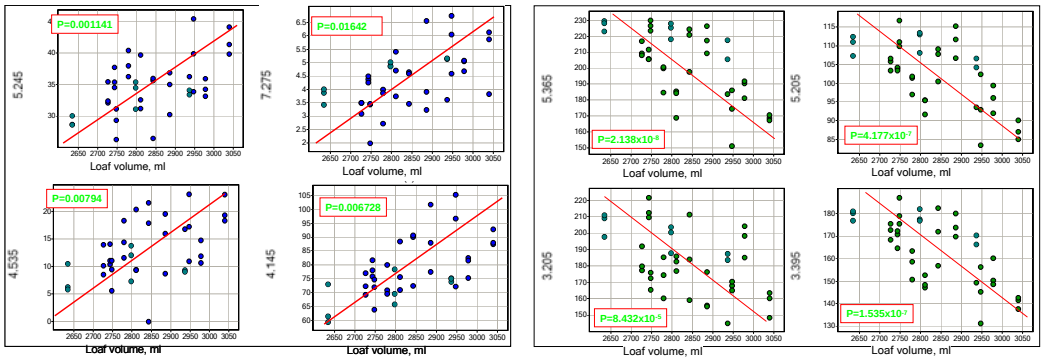
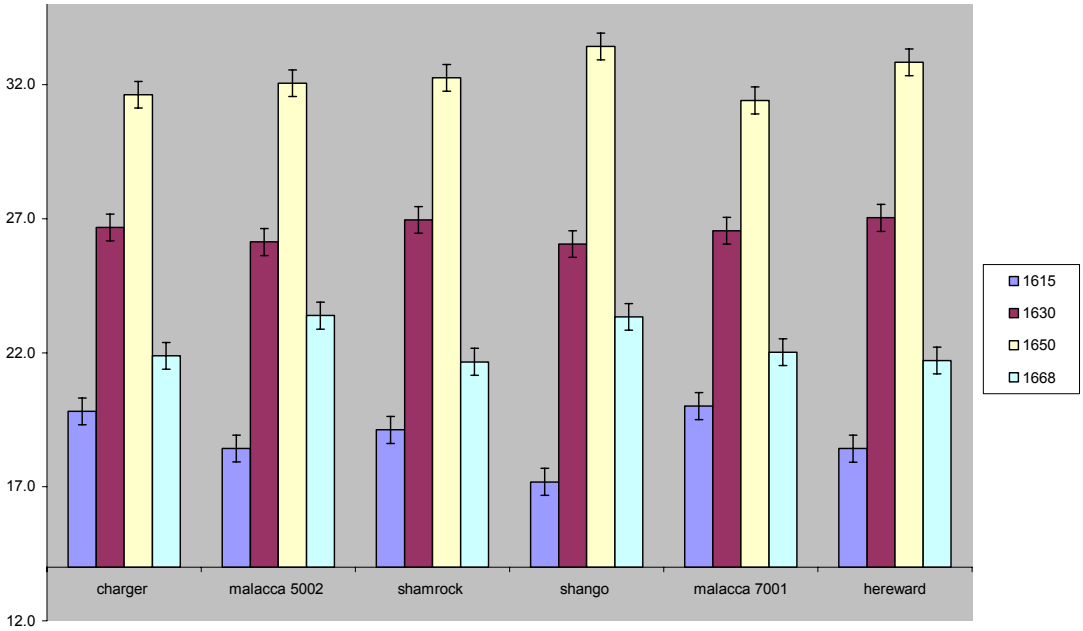


Figure 4. Relative intensities at four frequencies for fully hydrated flours. Error bars represent standard deviations of 5 replicates



New methods for assessment of fully hydrated flours using FT-IR were established within the project allowing comparisons to be made between both parental lines (Figure 4) and between lines selected to represent a range of performance for each product.

Although the new techniques were successfully applied within the project, the differences in FT-IR between samples having different properties were sufficiently small that no strong correlations could be identified. While it is recognised that FT-IR has been successfully applied to assess changes in dough during processing, it appears that the technique is less well-adapted at this stage for the quantification of differences between flours in respect of their baking performance.

Technical summary

- Three recombinant doubled haploid (DH) populations to determine the genetic control of end use quality were successfully developed.
- Methods for the objective assessment of final product quality for UK bread and puff pastry were developed and successfully implemented to generate a significant database of functionality data related to the doubled haploid populations.
- Significant new understanding of raw material functionality and processing attributes for the production of high value baked goods was successfully generated through the identification of new, statistically significant QTL which have been robust over 2 growing seasons.
- New relationships between processing quality and composition were established for flour quality tests and potential new opportunities for the biochemical techniques used were identified.

Conclusions

- The project has been a ‘once in a lifetime’ opportunity to make a step change in UK breeding for the development of high value wheat varieties.
- The project is at the forefront of similar work globally, with other initiatives smaller in scope and less-well resolved in terms of product assessment.
- The project has been a significant financial undertaking but represents good value due to the considerable additional efforts donated to the project by participants and the opportunity to give the UK grain chain a significant step forward in competitiveness.
- The outputs from the project have considerable value both in terms of short to mid-term breeding exploitation as well as through the resource created for future work.

Implications for levy payers

The UK wheat supply chain has seen significant changes in the last 40 years with a move from a market dependent on imports to one where the majority of wheat for a range of uses is home-grown. With increasing requirements and competitiveness in the marketplace, ongoing improvements in wheat yield, robustness and quality are fundamental to further improvements in the supply chain. This project gives plant breeders a new suite of opportunities for targeted breeding for quality and shows the potential for development of new varieties which have enhanced performance over those which are currently available. In addition, the work has shown new links between different aspects of processing quality and the genetic factors controlling them. It is anticipated that this will lead to further integration of the supply chain as the new knowledge will be used to better identify processing targets for breeding programmes and to give growers as well as breeders greater assurance in meeting them.

The project has represented a unique opportunity for representatives from across the supply chain to work together to generate improvements for all. By using end product quality as the target response, the project has provided wheat breeders with clearer targets for specific end uses, will lead to growers and processors having wheat better-adapted for their needs and the cereal science community with a major advance in the understanding of the genetic controls on wheat quality.

TECHNICAL REPORT

1. Introduction

The UK flour milling industry uses some 5.5M tonnes of wheat per year in the production of 4.4M tonnes of flour (**nabim**, 2007). The majority (85%) of the wheat used is UK-grown and this represents an approximate doubling in the usage of home-grown wheat over the last 30 years. The main utilisation of UK flour is in bread production with approximately 63% being used in the production of white, brown and wholemeal products. The supply chain is integrated with a strong emphasis on individual wheat varieties for the definition of contracts and specifications. Varieties are submitted for evaluation in respect of end use processing during the latter stages of their development and this is formally undertaken as part of National and Recommended List testing protocols. Individual varieties are placed in one of four **nabim** Groups as a result of these and other assessments, with the intention that the entire production chain is able to assess the likely value of a given variety for a particular application.

The substantial increase in the use of home-grown wheat in the last few decades has been driven by changes in trading arrangements with the development of the European Union but has been facilitated by representatives of various points within the supply chain. Plant breeders have focused on developing the required balance of quality characteristics for UK breadmaking processes as well as furthering the underpinning science on which this is based. One of the major advances occurred as a result of work undertaken at the former Plant Breeding Institute in which the relative contributions of different high molecular weight glutenin (HMWG) subunits were characterised (Payne, 1987; Payne *et al.*, 1987). Advances in agronomy allied to improvements in the genetics relating to functionality during baking have resulted in the current position where a substantial proportion of the UK wheat production results in moderately high protein content, hard milling wheat which may be used for a wide range of yeast leavened goods.

Changes in bread production methods have also facilitated the transition from net importation of wheat to a slight positive export balance. During the 1950s, workers at the former British Baking Industries Research Association (now part of CCFRA) developed a means of mixing and processing dough which removed the requirement for bulk fermentation of the mixed dough, allowing a significant reduction in time for

plant bread production. A short (2-5min) period of intense mixing (delivering 11Whkg^{-1}) allied to the use of an oxidising improver and high melting point fat was shown to produce a dough which could be divided immediately after mixing prior to a short intermediate proof, final moulding, final proof and baking. This significantly reduced the length of time needed for bread production as well as improving space efficiency in bakeries by dispensing with the need for dough holding areas (previously required for bulk fermentation). An added advantage of the use of CBP, however, was that flour having lower levels of protein could be used to produce bread. It is generally recognised that longer time processes tend to demand more from flour protein and a change to a shorter process allowed UK-grown wheats having lower protein contents than the North American or Australian wheats previously used in the 1950s and 1960s to form the main component of many breadmaking grists.

Although the current situation for wheat production and use in the UK owes much to these developments, it is recognised that further development of the present position is desirable but has a number of associated challenges. It is accepted that the true test of a given variety's 'quality' is its performance during milling and baking. Such testing requires significant quantities of seed, however, and so is not possible during the early stages of breeding. A number of smaller-scale and biochemical tests have been used historically to gain information at earlier stages but while these may be very successful for characteristics controlled by major genes, e.g. wheat hardness, they are less useful for the more complex traits associated with milling and baking performance. In addition, much of the selection at earlier generations is related to plant physiology and agronomic factors rather than those based on end use functionality. The resultant situation is one which contains significant risks for plant breeders where a realistic assessment of milling and baking performance is only likely after several years of development of a new variety. As a result there is a requirement for more robust methods of selection of early generation lines which include measures of required functionality as well as the full range of factors which influence agronomy. This project addresses this need by developing new markers and quantitative trait loci (QTL) associated with defined quality traits which may be used in marker-assisted breeding strategies to ensure that the appropriate selections are made.

Another aspect for further development in the UK market is that related to the 15% of imported wheat still required as a result of the particular properties required in some baked goods. It is reasoned that an improved understanding of the genetic basis for wheat and flour properties (such as protein content) which are known to impact on baking performance will facilitate UK breeding programmes aiming to produce further increases in baking quality and thus further reduce the need for imports.

As highlighted above, there have been many successes in developing the underpinning science linking wheat genetics to quality. Major genes controlling endosperm texture as well as the knowledge of the impact on processing of the different HMWGs have dramatically increased the tools available to plant breeders. Nevertheless, it is recognised that the wide range of properties which impact on milling and baking performance mean that the genetic control of such characters is likely to be based on effects from many minor genes. Although the understanding of HMWG functionality is extremely important, it is estimated that variation in these key components accounts for 30-50% of the total variation in dough mixing properties. This leaves a significant proportion of variability to be accounted for. In addition, dough mixing properties themselves would not be expected to account for all the product variation observed when producing baked goods. It is clear, therefore, that work to develop an understanding of the genetic controls for key bakery product attributes is a key priority for cereal science over the coming years. Indeed the tools available to breeders for selection based on markers and for the development of QTL are relatively new and so previous work in this area has been restricted in terms of assessment of baking performance, having been focused on indicators of performance such as hardness, protein content and dough properties.

An early study (Rousset *et al.*, 2001) developed links between wheat genetics and baking performance by assessing variation in group 1 chromosomes in relation to dough mixing time and loaf volume. Although this would be expected to further highlight the effects of HMWGs (encoded on the group 1 chromosomes), results also indicated that there were QTL influencing wheat protein content as well as loaf volume which were independent of these genes. The results of this study were published following the initial proposal development for the current project but it was recognised at this time that two other groups, one in Australia and one in France, were also

intending to work in this area. The output from these concurrent projects has recently been published and represents the most directly related background work.

The first publication is from the Australian group (Kuchel *et al.*, 2006) and details their work carried out using a single doubled haploid population based on two Australian varieties (Trident and Molineux). A number of QTL were identified which related to dough rheology, loaf volume and crumb score as well as milling yield, flour protein content and colour. Some of these where related to HMWG and low molecular weight glutenin (LMWG) loci on chromosomes 1A and 1B but a number of other QTL were found in alternative locations including 2A (dough strength and loaf volume), 3A (loaf volume and crumb quality), 6A (milling yield and flour protein content) and 7B (flour colour). Of these, the QTL on 3A is of interest as this confirms previous work in the UK which identified 3A as relevant in relation to protein properties and baking performance (Law *et al.*, 2005).

The French work was also undertaken using a single population (Renan × Récital) comprising 194 recombinant inbred lines grown in 3 environments (Groos *et al.*, 2007). Samples were assessed using a standard French breadmaking test which comprised both a scoring system based on subjective scores from trained assessors as well as determination of loaf volume. In addition, a number of quality parameters were collected including flour protein content, grain hardness and Alveograph parameters. While these were linked to the results obtained in the breadmaking evaluation, no combination of these parameters gave a good prediction of the values derived as part of the bread quality assessment. However, as with the Australian work, a number of novel QTL were identified for baking performance. These included 1A, 1B, 3A and 7B, all of which were also identified by Kuchel *et al.* (2006), as well as new QTL on 2B (dough and bread scores), 5B (loaf volume, bread and crumb scores), 6B (bread and crumb scores) and 7A (loaf volume and bread score). In common with the previous study, Groos *et al.* (2007) also found that the QTL for bread volume on chromosome 1A is not due to any of the known grain storage genes.

Although both these studies used a combination of subjective as well as objective methods of analysis, they clearly illustrate the potential merits of the approach adopted in the current project. The advantages of the work described in this report

are the use of three rather than single populations as well as the development of a wider range of objectively determined parameters to describe baked product quality.

1.1 Development of a wheat quality map

While it is generally accepted that test baking is the true test of a given wheat or flour's properties, there exist a wide range of wheat and flour quality tests which may be used to assess individual aspects of functionality. As a result, previous workers have attempted to use a wide range of quality assessment techniques in combination to develop ways of quality mapping wheat samples or predicting their likely performance. An extensive study by Andersson *et al.* (1994) showed that samples could be classified using principal components analysis (PCA) on the basis of wheat type (spring versus winter) as well as harvest year. While models were generated using a range of flour quality data for the prediction of loaf volume, these were heavily dependent on protein content which explained up to 47.5% of the variation in loaf volume alone. Subsequent work using UK data from National and Recommended List data for the development of NIR calibrations also saw the wheat and flour quality data being used to develop an equation to predict Chorleywood Bread Process (CBP) loaf volume (Millar, 2003). The equation had a squared correlation of 0.39 and a standard error of 161ml, neither of which would lend confidence in the predictive ability of such an approach. Multivariate data from individual instruments has also been used in the development of quality maps with the Reomixer being recently used for this purpose in an HGCA-funded project (Anderson, 2003). While the primary range of data used in this assessment was limited to that emanating from the instrument alone, other relevant parameters such as loaf volume were projected onto the quality maps thus derived. This showed that the data from the instrument could be used to define an area in principal component space within which good breadmaking varieties tended to reside. Given these previous studies, it is clear that the factors contributing to baking performance are complex and inter-connected such that a unifying relationship with other measured parameters is unlikely to exist. Nevertheless, as a means of demonstrating the merits of following a QTL approach to assessing wheat quality, a limited quality map assessment within the current study has been undertaken to show which of the currently measured wheat and flour parameters are relevant as well as giving a 'baseline' assessment of variability with respect to wheat variety and growing location.

1.2 Development of standard methods of bakery product assessment

While it was recognised that the wheat and flour assessment methods available for use in the project were based on established and previously validated standard methods, the methods for assessing end products were specific to each individual group. Given that a key element of the project was the use of objective methods of baked product analysis rather than a reliance on subjective, 'expert' scores, it was important to ensure that the measurement methods were being applied consistently and that any results would be comparable to those achieved by industry best practice. As a result, an exercise was undertaken for the methods used for bread analysis within which participating laboratories were sent loaves from common production batches for assessment. As the puff pastry assessment was undertaken by one group, they also developed new methods for assessing puff pastry and validated these internally. Again this work was based on the premise that new, objective methods were to be applied which would result in data being captured which represented the differences which would be taken into account by an experienced human assessor.

Subsequent to the ring trial and as a means of supporting ongoing development within each group on the means of determining loaf volume, a standard set of reference loaves was circulated to the group interspersed with regular determination of their reference volume by water displacement. This was used as a final check of the consistency of response for each laboratory in advance of the main body of work commencing and in some cases was used to make a final decision on the method to be used for volume assessment.

1.2.1 Development of a specific puff pastry method

It is thought that this project represents the first time that puff pastry has been assessed on such a major scale and as a result there was a need to establish an efficient protocol for producing and assessing the samples. In traditional assessment of puff pastry by expert bakers, an overall quality score for the finished product is generally given. This typically combines parameters such as the height and internal structure of the pastry. However, the scope of the current project necessitated the development of quality parameters which were more objective. This process was complicated by the complex nature of puff pastry in terms of production, assessment and the intrinsically variable nature of the product. To ensure that the approach used in the project was optimised, a total of 8 potential methods were evaluated. These

methods represented various interpretations of the two main categories of puff pastry production:

- Long time, including long rests between each processing step or
- Short time, with shorter rests between stages.

Shorter time processes favour flours with a lower protein content or weaker protein quality. Longer processes tend to favour higher protein content flours, with the extra rest time allowing for greater relaxation of the gluten and better performance in general.

1.2.2 Genetic map development and QTL analysis

The location of genes controlling complex traits, such as components of wheat processing quality which are measured quantitatively, is now possible by the application of marker-mediated genetic analysis for QTL analysis of defined crosses. This relies on the establishment of an association between the segregation of known marker alleles with differences in phenotypic expression of the trait.
Marker-mediated genetic analysis has four steps:

1) The development of an appropriate recombinant population between parents differing in phenotype for the character(s) of interest. In the present series of experiments recombinant doubled haploid populations developed using the maize cross system were developed by the Breeders and JIC from F_1s of the three carefully chosen crosses, Malacca x Charger, Hereward x Malacca and Shamrock x Shango differing in quality attributes.
2) A genetic map of the recombinant population is developed by characterizing all lines for their allelic constitution at molecular marker loci dispersed throughout the genome and identified as being polymorphic in the cross from examination of the parental lines.
3) Concurrent with map development is the evaluation of seed from the recombinant lines in appropriate replicated and randomised experiments for the traits of interest.
4) The final step is to use statistical procedures to combine the map and trait data so as to partition the genotypic variation into effects of individual QTL and to locate these relative to the marker loci in defined chromosome segments.

1.3 Metabolomics

Metabolomics is the study of small-molecule metabolite profiles. The metabolome represents the collection of all metabolites in a given biological organism, which are the end products of its gene expression. Unlike mRNA gene expression data and proteomic analyses, metabolic profiling can give an instantaneous snapshot of the physiology of the biological system.

Metabolomics relies on the collection and analysis of analytical datasets and the mining of that data, often using multivariate statistics, to integrate the analytical results with biochemical pathway information. Due to the large number of samples contained within a typical metabolomics experiment, the requirement for robust high-throughput methods to generate comparable datasets is a key requirement.

In general, NMR spectroscopy and Mass Spectrometry techniques are most widely used as they tend to provide a comprehensive coverage of the metabolome, and are especially useful in identifying key changes in the primary metabolome. In this project, high throughput NMR screening has been employed. NMR is a non-destructive technique, and unlike mass spectrometry is absolutely quantitative. Although the technique is less sensitive than mass spectrometry, it is non-selective and the method returns an information-rich fingerprint of many overlapping peaks. These peaks result from the hydrogen signals of every metabolite which is soluble in the chosen solvent system.

In terms of utility of the dataset, multivariate methods such as PCA are extremely useful in gaining a broad overview of the NMR datasets. Information explaining separations of clusters within this type of analysis can easily be extracted and represent signatures of the metabolites responsible for the differences between samples.

The above approach has been used in this project and has been supplemented with classical univariate statistical approaches to explore correlations within the dataset and to map metabolite signals onto breadmaking quality data.

1.4 Fourier transform infrared spectroscopy

Fourier transform infrared spectroscopy (FT-IR) is a very versatile method for examining the secondary structure of proteins and has been used to considerable effect in the study of gluten proteins (Almutawah *et al.*, 2007; Georget and Belton, 2006; Wellner *et al.*, 2006). However early attempts to examine different samples of gluten from parents and crosses were not successful. This appears to be due to two factors. The first is simply the biological variation of samples selected from different parts of the crop and the second is the variation in spectra that are observed simply due to differential handling of the gluten samples. Previous work (Georget and Belton, 2006; Wellner *et al.*, 2006; Wellner *et al.*, 2005) has shown that mechanical handling of gluten samples can make substantial differences to the observed spectra. Prior to undertaking work in the project, therefore, a new approach for sample preparation needed to be developed.

1.5 Project rationale and objectives

The need to improve selection for end use quality led to discussions between commercial plant breeders, academic researchers and end users in November 2000. At the meeting it was proposed that a LINK research programme should be developed to investigate wheat functionality through breeding and end use. Although great strides had been made during the previous two decades, it was concluded that 60% of the variation in end use quality between varieties was not satisfactorily explained by predictive tests currently available to breeders. Early generation selection for quality was thus compromised and end user value was only being satisfactorily assessed by test generation of final products after selection was essentially fixed. It was decided that 3 random inbred populations of breeding lines from key parents should be analyzed by end users and genetic mapping techniques employed to identify key quality QTL.

As a result of the discussions which led to the initiation of the project, the following objectives for the project were agreed.

- To improve the understanding of raw material functionality and processing (primarily in relation to end use qualities of milling and breadmaking) by creating a benchmark “quality map” using current breadmaking wheat varieties and a combination of existing data plus novel quality parameters.

- To facilitate matching end users' requirements to appropriate raw material through targeted wheat breeding via the development of genetic maps of genes for processing quality.

- To evaluate experimental populations of doubled haploid lines produced from established parent varieties representing "key" variation in UK winter wheat germplasm for processing quality traits.

2. Materials and methods

2.1 Population development

During early discussions it became clear that sufficient variation in key quality characters existed in current varieties for useful populations to be produced. It was decided that irrespective of the presence of interesting quality characters, no populations would be made with varieties containing the 1B/1R rye translocation (found in many feed wheats and some quality wheat varieties) or with varieties with soft endosperm. A provisional list of varieties of interest to end users and breeders was drawn up: Hereward, Malacca, Charger, Xi19, Shamrock, Shango, Soissons, Buster and Spark. Xi19, Soissons and Spark were discarded due to variable end use performance, dwarfing gene incompatibility and undesirable growth habit segregation.

Three populations of interest to breeders and end users were agreed: Buster × Hereward, Malacca × Charger and Shango × Shamrock. These populations would segregate for a number of intractable quality characters such as flour colour, brightness, extraction rate, dough performance and final bake performance. Concern was expressed that high molecular weight (HMW) glutenins were also in segregation in two of the populations but it was felt that statistical analysis would allow adjustment for any end user quality effects due to this segregation.

Doubled haploid populations were produced using the inter-specific wheat × maize hybridization methodology first developed at The Plant Breeding Institute and modified by the John Innes Centre and in-house at Syngenta and RAGT. The John Innes Centre was sub-contracted to produce doubled haploid lines for Nickerson. The integrity of the doubled haploid populations was preliminarily assessed by breeders using HMW glutenin sub-unit analysis.

Doubled haploid populations were scheduled to be developed from 2001 until August 2003 when 600g of each doubled haploid line was to be available for first field production in the 2003 – 2004 season. Doubled haploid population development within this time frame proved very difficult. During 2003 it became apparent that there would be insufficient lines produced from the Buster × Hereward cross to merit further analysis (a total of only 54 lines were produced) and only 84 lines from the Shango × Shamrock cross would be available, a number that was deemed just sufficient for the proposed research. As a consequence RAGT offered a doubled haploid population from the cross Hereward × Malacca for use by the research consortium. This offer was accepted and subsequently bulked grain of doubled haploid lines from the crosses Malacca × Charger, Hereward × Malacca and Shamrock × Shango were made available to end users.

Due to the problems in doubled haploid production and the poor harvest conditions in 2004 leading to unacceptable Hagberg Falling Numbers for end user testing, the bulk grain production from the 2004 harvest was compromised. Grain production in 2005 and 2006 was more successful. By spring 2004, 120 doubled haploid lines from the Malacca × Charger cross, 120 lines from the Hereward × Malacca cross and 87 lines from the Shamrock × Shango cross were available for field production. Bulk grain was produced of these lines from harvest 2005 and 2006.

2.2 Development of a wheat quality map

Samples were selected from the Recommended List (RL) trials in 2001 from 5 sites: Cambridge, Headley Hall, Wye, Cockle Park and Morley. The varieties included were Malacca, Rialto, Charger, Chatsworth, Fender, Hereward, Macro, Option, Phlebas, Shamrock, Soissons, Solstice and Xi19. Each sample was Bühler milled to give white flour and the flour was baked following both CBP and no time dough (Spiral) procedures. Assessment of the wheat, flour and bread was carried out using the following methods.

- Wheat specific weight - FTWG 20 (CCFRA, 2002).
- Grain length, breadth, perimeter and area by image analysis.
- Perten single kernel characterisation system (SKCS) following manufacturer's instructions.
- Grain moisture and protein content by wholegrain NIR.

- Bühler milling yield (extraction rate).
- Flour particle size distribution using a 75µm sieve.
- Flour moisture and protein content using a Perten 8611 NIR instrument - FTWG 14 (CCFRA, 2002).
- Damaged starch by Farrand - FTWG 05 (CCFRA, 2002).
- Flour Hagberg Falling Number - FTWG 06 (CCFRA, 2002).
- Flour Grade Colour - FTWG 07/4 (CCFRA, 2002).
- Branscan.
- Full Farinograph analysis of dough properties - FTWG 04 (CCFRA, 2002).
- Flour pentosan content.
- Endogenous cereal *alpha*-amylase activity by Ceralpha - FTWG 18 (CCFRA, 2002).
- Analysis of the gel protein fraction (Alava *et al.*, 2001).
- CBP and Spiral loaf volume by seed displacement.
- CBP bread crumb structure - baker's score.
- Image analysis of CBP bread crumb.

Datasets representing each of the materials assessed (wheat, flour and bread) were assessed using PCA to assess clustering of varieties/sites and partial least squares (PLS) regression to predict baking performance from grain and flour properties. All statistical analysis was carried out using Minitab 15.

2.3 Development of standard methods of bakery product assessment

2.3.1 Bread ring trial

Bread was produced to cover a range of volumes and internal structural attributes and also to produce representative loaves for both white and wholemeal production methods. To achieve this, four different commercially milled flour types were sourced to represent white bread flour, white general purpose flour, wholemeal bread flour and wholemeal flour from soft milling wheats. From these, bread was made using both CBP and no time dough (Spiral) procedures for both wheat flours but only CBP methods for both wholemeal flours. Thus a total of 6 baking variations were undertaken.

Samples produced under CBP conditions were made using a Tweedy 70 mixer operated under a pressure/vacuum regime of 1.5bar/0.34bar with the changeover

occurring after delivery of half of the required work input (11Wh/kg). No time doughs were produced using a Spiral mixer operated at 2min slow speed followed by 10min high speed. The formulations used were as given in Table 1.

Table 1. Formulation for bread production

	CBP and no time dough white bread (% on flour weight)	CBP wholemeal bread (% on flour weight)
Flour	100	100
Water	To Farinograph 600 Line	To Farinograph 600 Line
Yeast	2.5	2.5
Salt	2	2
Fat	1	2
Improver	1	2

The fat (Quartz, Pura Foods) and improver (Mistral, Cereform) had previously been selected for use throughout the remainder of the project.

Following dough production to a target final dough temperature of 30.5±1ºC, doughs were rounded using a conical moulder prior to a period of 7min intermediate proof. Dough pieces were then moulded (four piece for white and single piece for wholemeal) and proved for 50min prior to baking for 30min at 244ºC. The CBP white bread was produced with lids on the bread pans; the other products were all unlidded. Following cooling, 8 loaves were from each product were transported to each of the other laboratories participating in the ring trial.

On receipt, end users were asked to assess loaves for volume using standard seed displacement (and ranging techniques where available), crumb whiteness for white loaves (Tristimulus Y), yellowness for wholemeal (CIE b), crumb firmness by texture profile analysis (TPA) and crumb structure by C-Cell. Four laboratories participated in the exercise with Laboratory 4 submitting results for crumb firmness and C-Cell alone.

Loaf volume measurements were carried out using an in-house ranging instrument by Laboratory 1 and seed displacement for Laboratories 2 and 3. For crumb firmness assessment, users were asked to perform TPA measurements on discs of breadcrumb

cut from two loaves at each time point. A minimum of six measurements were to be made on each loaf at each time point. Measurements were made on Day 1 and Day 3 with bread being stored at 21^{o}C between assessments. For C-Cell, attention was required when slicing to ensure consistency between users. For 4-pieced bread, loaves were to be sliced using a knife in the centre, i.e. between pieces 2 and 3. The blade depth on the Graef C-Cell slicer was to be set at 15mm and, for each half loaf, slices were to be removed until the slices at the centre of pieces 2 and 3 were reached. These slices were to be used for C-Cell assessment. For single-pieced bread, the loaf was to be sliced in the centre using a knife and two 15mm slices cut using the blade slicer from each half. The second slice was to be used for C-Cell measurements.

Results for each laboratory were then collated and used to assess the consistency of results between participants.

2.3.2 Loaf volume calibration

Eight polystyrene loaves were obtained from TexVol Instruments AB (Viken, Sweden), the volumes of which had been specified to cover the entire range of both 400g and 800g unlidded loaves (4 loaves for each nominal size range). Reference volumes for each of these were determined using water displacement. A stable Micro Systems (SMS) TA-XT2 instrument was fitted with a circular attachment which was sufficiently large to hold each of the polystyrene loaves under water. A trough was placed on the base of the SMS instrument and filled with water at 20±0.5ºC, the density of which was determined using a hydrometer. The weight of the circular attachment was recorded prior to fitting to the instrument. To take the baseline reading, the circular attachment was submerged to a fixed depth using the instrument. When the water had settled, the live force reading from the instrument was reset and the circular attachment withdrawn from the water. Following drying of the surface of the circular attachment, the reading then given was recorded. The process was then repeated using one of the pre-weighed polystyrene loaves placed under the circular attachment. To ensure that any holes generated in the loaf ends as a result of placement on ranging instruments did not affect the results, a fixed weight of Blutack was also attached to the loaf ends. This Blutack was also attached to the circular attachment for the baseline reading.

The reference volume (ml) for each polystyrene loaf was then calculated as:

$$V = (F - F_0 + W + w) / \rho$$

where

F_0 = force reading (grams) on SMS after test with Blutack and circular attachment

F = force reading (grams) on SMS after test with polystyrene loaf, Blutack and circular attachment

W = weight of polystyrene loaf (grams)

w = weight of Blutack

ρ = density of water (grams per ml)

Polystyrene loaves were supplied to Laboratories 1, 2 and 3 for assessment using their own calibrated methods for determining loaf volume. Loaf volumes were assessed at CCFRA by water immersion prior to supply to each participant and assessed again in the same way on the loaves' return.

All participants had operational seed displacement techniques for assessing volume as well as ranging methods. Laboratory 1 used an in-house ranging system while Laboratories 2 and 3 both used TexVol instruments in laser and ultrasound configurations respectively.

2.3.3 Puff pastry method

While each of the two approaches for puff pastry production (long time and short time) have merits in commercial usage, it was recognised that the large number of samples to be assessed within this project placed some practical constraints on operators. To run a long time process effectively, significant overlaps in processing between samples would be required which, in turn, would create vulnerability to equipment breakdowns and power failures as well as posing difficulties for individual scheduling. As a result, a shorter time process was seen as being preferable and so the method used was based on such principles. Two bakers were trained in the method and statistical process control limits were generated using a control chart.

These limits were then used subsequently to validate each day's production within the project.

For objective assessment of puff pastry, it was recognised at an early stage in project planning that the standard measurements generated by C-Cell were not all relevant for assessment of the structure of a laminated product. Development work was then undertaken from which a revised list of the key parameters to be used for puff pastry assessment was developed. These were as follows:

- Average Height (pixels)
- Breadth (pixels)
- Coarse /Fine Clustering
- Number of Cells
- Cell Volume
- Average Cell Elongation
- Cell Alignment
- Non-Uniformity
- Wall Thickness (pixels)
- Count >10
- Max Bubble (pixels)

From these parameters a calculated score was developed following the results of an assessment of 52 pastry samples covering a range of properties. The scoring method was then validated using a further 16 samples assessed using C-Cell as well as by 8 expert assessors. A further 16 samples were then used as a final check on the reliability of the objective scoring method.

In addition to the C-Cell images of the internal pastry structure, further analysis was also carried out to assess the important aspects of product texture and shrinkage during baking. To generate these data the tenderness and layer count were measured using a specific texture analyser method and the dimensions of the pastry measured by hand using a calibrated set square.

These assessments resulted in the derivation of a further 6 quality parameters:

- Tenderness
- Layer Count
- Height (mm)
- Width (mm)
- Depth (mm)
- Calculated % Shrinkage

2.4 Grain production and supply

The three breeders involved in the project (Nickerson-Advanta Ltd, RAGT Seeds Ltd and Syngenta Seeds Ltd) produced doubled haploid populations of approximately 150 individuals from each cross and sufficient bulk grain (25 kg) of each line from two consecutive harvests to allow two years of end user testing. The production of grain from each doubled haploid line was carried out at two sites in each production year. Grain of each doubled haploid line from both sites was cleaned and pooled to produce a single, sound sample for end user testing. Although problems at the Nickerson site in 2005 led to less grain of the Hereward x Malacca population being produced than desired, sufficient grain was available for end user testing when the Nickerson sample was pooled with the sample from the RAGT site. The 2006 grain production went smoothly with no problems with the quantity of grain produced.

Grain production trials were grown following British Society of Plant Breeders (BSPB) National List (NL) trial protocols with nitrogen fertilisation adjusted to produce protein content acceptable to the end users. Trials were grown in three replications, randomised using appropriate experimental designs. Each doubled haploid line was grown in drilled plots, generally 6m x 1m, to simulate performance in farmers' fields. Bulk grain of the parents of each population was produced alongside each population. The breeders also carried out the cleaning, bulking and dispatch of grain samples to the end users.

The soundness of the harvest quality of the bulk grain samples was assessed by the use of NIR and Hagberg Falling Number tests of the parental controls growing alongside each population. In addition to the production of sound samples of grain for

the end user testing, basic field data, including plant height, ear emergence and yield, were collected by the breeders for each of the DH lines.

2.5 Sample handling for milling and baking

Wheat samples were supplied to each of the 3 organisations undertaking the milling part of sample assessment, ADM Milling, Allied Technical Centre and CCFRA. Each organisation was responsible for one population. Samples were supplied from two growing locations which were blended prior to milling and baking. In both years, samples were blended in a 50:50 ratio by weight for each of the two growing locations for that population. Where the quantities differed such that this split wasn't possible, the entirety of the smaller sample was blended with sufficient of the larger sample to ensure that the total quantity required for milling was retained. For the first year assessment (2005 harvest), a total sample mass of 25kg was used. Experience from this year and some changes in assessment led to the target quantity for 2006 harvest material being reduced to 22kg.

In each year, samples from each population were assessed in terms of the levels of non-wheat material and where necessary, sample cleaning was undertaken following the local procedures of the site at which each assessment had been carried out. In each case the system used was capable of removing chaff, stones, mudballs, weed seeds and broken grain. Prior to milling, samples were stored at a moisture content of <15% and within a temperature range of 10-15ºC.

Following milling, flour samples were supplied for the production of the 4 bakery products selected for evaluation within the project: CBP white bread, CBP wholemeal bread, no time dough (Spiral) white bread and puff pastry. Flour samples were blended to give three white flour samples of the appropriate sizes for generation of the products using white flour. Wholemeal flour was produced using white flour along with the bran and offal fractions in the required proportions. A sub-sample of each of the finished white and wholemeal flours was then used for flour assessment and the data thus generated, along with the samples, were supplied to the organisation undertaking the relevant end product generation and assessment. Samples were produced, dispatched and stored in a regular way. Wholemeal flours were baked within 3 weeks of milling to avoid changes associated with longer storage having an impact on the results.

Following production and assessment of each of the products, the data generated during milling, flour assessment and baking phases were collated by each of the organisations responsible. These data were then reviewed by the combined end user group prior to supplying the data for QTL analysis.

2.5.1 Flour milling

All milling was performed using Bühler MLU 202 laboratory mills using an agreed common procedure. Samples were conditioned prior to milling by assessing the moisture content of the starting grain and adding the required amount of water to give a grain moisture content of 16%. Grain was held for at least 16h prior to milling and the moisture content was checked to be in the range 15-16% prior to feeding onto the mill.

Mills were operated in a controlled environment within a temperature range of 18-22ºC and a relative humidity of 60-70%. The roll gaps and cover sizes were standardised in accordance with Table 2.

Table 2. Operating parameters for Bühler mills

Roll gaps (mm)		**Sifter covers (µm)**			
			1	2	3
First break	0.06	Break (top)	710	600	530
Third break	0.04	Break (bottom)	140	140	140
First reduction	0.03	Reduction (top)	180	180	140
Third reduction	0.02	Reduction (bottom)	180	180	140

To produce white flour, the bran and offal fractions were each passed twice through an impact finisher. All flour fractions were then given a final sieving prior to blending for 20min. Wholemeal flour was produced by taking all the flour fractions and adding the bran offal in the correct proportions prior to blending.

Milling performance was estimated by calculating white flour extraction rate (feed basis) as well as the proportion of flour emanating from bran and offal finishing.

2.5.2 Wheat and flour analysis

All wheats were assessed for protein and moisture content as well as Hardness Index by NIR or near infrared transmittance (NIT). The calibrations used were based on protein content by Dumas, moisture content by oven drying and Hardness Index by Perten Single Kernel Characterisation System (SKCS). Similarly, flour protein and moisture contents were assessed by NIR with calibrations based on Dumas and oven moisture respectively. White flours were assessed by full Farinograph in accordance with Flour Testing Working Group Method FTWG 04 (CCFRA, 2002). Farinograph assessment for wholemeal flour was restricted to the determination of water absorption alone for baking. Both white and wholemeal flours were assessed for Hagberg Falling Number in accordance with FTWG 06 (CCFRA, 2002).

2.5.3 Bakery processing

Following the method development phase of the project, each end user organisation used validated local methods for baked production generation and assessment. For the bread products, the formulations and ingredients are detailed in Table 1. For puff pastry, the formulation used is given in Table 3.

Table 3. Formulation used for puff pastry production

Ingredient	% on flour weight
Flour	100
Pastry margarine	48
Water (2-6ºC)	46
Cake margarine	7
Salt	1.5

Samples of each baked product were assessed using previously agreed objective techniques. For bread, loaves were submitted to loaf volume determination, crumb colour (CIE L*a*b* and Tristimulus X, Y, Z), crumb firmness by texture profile analysis on Days 1, 3 and 7 (Day 7 for 2005 harvest only) and C-Cell. Crumb firmness was not assessed on Day 7 for the 2006 harvest samples due to problems experienced in storage during the analysis of 2005 material.

The specific details of each of the local methods used by the four organisations undertaking bakery assessment are included in Appendix 1.

2.5.4 Analysis of milling and baking data

Following analysis of all samples in both harvest years, each end user group collated all the results generated in spreadsheet format for further analysis. The method development for puff pastry resulted in a subset of C-Cell parameters being used. For the bread analysis, all standard C-Cell parameters were collected for potential submission for QTL analysis. The results from the 2005 harvest were used, however, to generate a reduced set of parameters from texture and C-Cell analysis, some of which were calculated from the basic output. These were then used for treatment of the data from both years and were the focus for the final QTL analysis. The final reduced set of parameters was as follows:

- Loaf volume
- L*
- Firmness Day 1
- Firmness Day 3
- Firmness Day 3/loaf volume
- Firmness Day 3 - Firmness Day 1
- Firmness Day 3 - Firmness Day 1/loaf volume
- The sum of left and right concavity
- The sum of left and right concavity/loaf volume
- The sum of left and right concavity/maximum slice height
- Number of cells
- Number of cells/slice area
- Cell wall thickness
- Cell diameter
- Coarse cell volume

In addition, a number of specific measurements of key dough parameters were also collected for the wholemeal product:

- Dough consistency ex. mixer
- Dough consistency 10 minutes after completion of mixing

- Degree of softening
- Mix time
- Dough temperature ex. mixer
- Dough temperature 10 minutes after completion of mixing
- Gross water to give a consistency of 410 Brabender Units
- Manual stickiness score

These parameters were supplied for QTL analysis in addition to the data generated from the milling and wheat/flour quality testing. In addition, samples were selected from each population/product combination which were representative of the poorest, the average and the best performing sample. For the bread products, these samples were assessed using a sub-set of the parameters assessed:

- Loaf volume
- L*
- Firmness Day 3
- Number of cells/slice area
- Slice brightness
- Cell wall thickness
- Cell diameter

Each sample was number ranked for each of the parameters above, i.e. for a population of 100 lines, bread samples were ranked from 1-100 for loaf volume, L* etc. The rankings for all parameters were then added up to give an overall ranking. The samples giving the best and worst overall ranking were selected as well as a sample representing the midpoint.

For pastry products, product selection was undertaken by an expert panel of assessors using visual examination of C-Cell images as well as all the calculated score.

The samples which were selected in each harvest year were then used as the basis for sample selection for further analysis by metabolomics or FT-IR.

2.6 The development of the genetic maps of the three wheat crosses

Simple Sequence Repeats (SSR - also known as microsatellites) molecular markers were primarily used for mapping (Somers *et al.*, 2004), supplemented by Diversity Array Technology (DArT) markers (to fill any gaps) (Jaccoud *et al.*, 2001). Grain storage protein markers, particularly variation in the HMWGs identified through SDS-PAGE, were also used in map development where they were polymorphic.

2.6.1 SSR methods

DNA was extracted from the leaves of seedlings germinated in Petri dishes for the parents and each of the individual doubled haploid recombinant lines of the three crosses, as described in Snape *et al.* (2007).

Various sets of SSR marker primers are publicly available (see 'GrainGenes' **http://wheat.pw.usda.gov)**. The JIC has established a core set of around 1000 primers from these different sources, characterised for ease of use, polymorphism levels, and map location and distribution. These were applied initially to the parents to identify the polymorphic primers appropriate to each cross. Since the same sets of markers were used for assessing polymorphisms in each population, this would also allow cross-referencing of markers and QTL maps for comparative mapping purposes. Following the identification of a set of polymorphic markers for each cross, chosen to cover the genome as far as possible, each line of each of the three populations was genotyped for all available polymorphic markers.

SSRs were analysed as described in Roder *et al.* (1995), being run on 5% polyacrylamide gels and visualised by silver staining (Bassam *et al.*, 1991).

2.6.2 DArT technology

This is a new DNA-chip-based technology where DNA of each individual line is hybridized to a chip containing around 1000 DNA fragments, and polymorphisms are revealed by the presence or absence of hybridization of genomic DNA from each individual line. DArT is able to classify germplasm without previous knowledge of genomic sequences and allows several hundred polymorphisms to be simultaneously indentified. DArT in wheat is carried out as a service by a specialist company, Triticarte Pty Ltd, Australia (www.triticarte.com.au/) and DNA of the parents and each individual recombinant doubled haploid lines was sent to Australia for analysis.

Genotyping data were received back as a +/- polymorphism for each marker for each line.

2.6.3 Map development

All genotyping data from the different marker procedures on each population were entered into Excel spreadsheets. These data were then processed to develop the genetic map for each population using JoinMap software.

2.7 The location of new QTL for end use quality

2.7.1 Collection of the phenotypic data

In addition to genotype mapping, accurate QTL analysis also relies on precise phenotyping. For quality analysis, this requires seed from trials grown under an agronomy system appropriate for a baking sample. Seed for processing was obtained from pooled samples in each harvest year, 2005 and 2006, as described in Section 2.4.

2.7.2 Data collation

Data from the end users, for each of the products, in each year, for individual lines within each cross, were supplied to the JIC as Excel spreadsheets. The data obtained from the different end users is shown in Table 4.

Table 4. Distribution of processing duties to end users and sources of data, respectively

			Cross	
		M×C	**H×M**	**S×S**
	Milling	ATC	CCFRA	ADM
Product	**Wholemeal**	RHMT	RHMT	RHMT
	CBP White	ATC	ATC	ATC
	Spiral White	CCFRA	CCFRA	CCFRA
	Puff Pastry	ADM	ADM	ADM

The details of the traits analysed for the different products on each harvest sample, 2005 and 2006, are shown in Tables 5-9. In 2006, some new derived parameters were produced by the end users for analysis, and these were retrospectively added to the 2005 data set.

Table 5. Milling traits data supplied by end users for the 2005 and 2006 harvests

	M×C Milling data 2006	2006	2005
Wheat	Wheat moisture content, %	**1**	**1**
	Wheat protein content, % (Nx5.7) dmb	**2**	**2**
	Hardness Index (NIR),	**3**	**3**
	Straight run (white), %	**4**	**4**
Milling	Bran finisher flour, %	**5**	**5**
	Offal finisher flour, %	**6**	**6**
	Protein content, % (N x5.7) @14%mc	**7**	**7**
White Flour	Protein content, % (N x5.7) dmb	**8**	
	Water absorption (600 line), %	**9**	**8**
	Water absorption (600 line), % (@14% moisture)	**10**	
	Development time, min	**11**	**9**
	Stability , min.	**12**	**10**
	Degree of softening, BU	**13**	**11**
	Falling No.(s)	**14**	**12**
	Water absorption (600 line), %	**15**	**13**
Wholemeal flour	Falling No.(on KT- ground wholemeal), s	**16**	**15**
	Water absorption (600 line), % (@14% moisture)	**17**	**14**

Table 5. (contd) Milling traits data supplied by end users for the 2005 and 2006 harvests

	H×M milling traits 2006	2006	2005
Wheat	Wheat moisture content, % NIR	**1**	**1**
	Wheat protein content, % (Nx5.7) dmb NIR	**2**	**2**
	Hardness Index (NIR),	**3**	**3**
Milling	Straight run extraction rate (white), %	**4**	**4**
	Bran finisher flour, %	**5**	**5**
	Offal finisher flour, %	**6**	**6**
White Flour	Moisture content, %	**7**	**7**
	Protein content, % (N x5.7) @14%mc	**8**	**8**
	Water absorption (600 line), %	**9**	**9**
	Water absorption 14%mb	**10**	**10**
	Development time, min	**11**	**11**
	Stability, min.	**12**	**12**
	Degree of softening, BU	**13**	**13**
	Falling No.(s)	**14**	**14**
Wholemeal flour	Water absorption (600 line), %	**15**	**15**
	Falling No.	**16**	**16**
White flour	Protein content, % (N x5.7) dmb	**17**	
	Protein loss	**18**	

Table 5. (contd) Milling traits data supplied by end users for the 2005 and 2006 harvests

	S×S Milling Traits 2006	2006	2005
Whole Grain Foss Results	Protein	**1**	**1**
	Moisture	**2**	**2**
	SKCS	**3**	**3**
White Flour Lab Results	Mill Recovery	**4**	**4**
	Extraction Rate	**5**	**5**
	Protein loss on Milling	**6**	
	Protein (as is)	**7**	
	Protein (DMB)	**8**	
	Hagberg	**9**	**6**
	WAB	**10**	**7**
	DDT	**11**	**8**
	Stability	**12**	**9**
	Deg of Softening	**13**	**10**
Wholemeal Lab Results	Hagberg	**14**	**11**
	WAB	**15**	**12**

Table 6. Puff pastry trait data supplied by end users for the 2005 and 2006 harvests

2005 and 2006
Avg Height (avg)/px
Avg No. of Cells
Avg Cell Volume
Avg Count >10
Avg Calc Score
Avg Tenderness
Avg Breadth / px
Avg Coarse /Fine Clustering
Avg Cell Alignment
Avg Non-Uniformity
Avg Wall Thickness / px
Avg Layer Count
Avg Height (mm)
Avg Width (mm)
Avg Depth (mm)
Avg Calc % Shrinkage

Table 7. CBP white trait data supplied by end users for the 2005 and 2006 harvests

CBP White Traits	Code 2006	Code 2005	Reduced data set 2006
Loaf volume, ml	**1**	**1**	**1**
L*	**2**	**2**	**2**
a*	**3**	**3**	
b*	**4**	**4**	
X	**5**	**5**	
Y	**6**	**6**	
Z	**7**	**7**	
Firmness @25% compression (Day 1), N	**8**	**8**	**3**
Firmness @25% compression (Day 3), N	**9**	**9**	**4**
Firmness (DOP3)/loaf vol	**10**		**5**
Firmness @25% compression (Day 3-1), N	**11**		**6**
Firmness @25% compression (Day 3-1)/loaf vol	**12**		**7**
Slice_Area	**13**	**10**	
Average Height_Max	**14**	**11**	**8**
Height_Avg	**15**	**12**	
Breadth	**16**	**13**	
Height_ Breadth	**17**	**14**	
Wrapper_Length	**18**	**15**	
Total_ Concavity	**19**	**16**	
Left_ Concavity	**20**	**17**	
Right_ Concavity	**21**	**18**	
Top_ Concavity	**22**	**19**	
left+right concavity	**23**		**9**
left+right concavity/loaf vol	**24**		**10**
left+right concavity/max slice height	**25**		**11**
Bottom_ Concavity	**26**	**20**	
Left_Bk_Area	**27**	**21**	
Right_Bk_Area	**28**	**22**	
Left_Bk_Height	**29**	**23**	
Right_Bk_Height	**30**	**24**	
Left_Bk_Depth	**31**	**25**	
Right_Bk_Depth	**32**	**26**	
Left_Bk_Pos	**33**	**27**	

Table 7. (contd) CBP white trait data supplied by end users for the 2005 and 2006 harvests

CBP White Traits	Code 2006	Code 2005	Reduced data set 2006
Right_Bk_Pos	**34**	**28**	
Top_Left_Shoulder	**35**	**29**	
Top_Right_Shoulder	**36**	**30**	
Bottom_ Left_Roundness	**36**	**31**	
Bottom_ Right_Roundness	**37**	**32**	
Slice_ Brightness	**38**	**33**	
Cell_ Contrast	**40**	**34**	
Number_ of_Cells	**41**	**34**	**12**
Cells/unit area	**42**		**13**
Number_ of_Holes	**43**	**36**	
Area_of_ Cells	**44**	**36**	
Area_of_ Holes	**45**	**37**	
Volume_ of_Holes	**46**	**38**	
Wall_ Thickness	**47**	**40**	**14**
Cell_ Diameter	**48**	**41**	**15**
Cell_Vol_Range	**49**	**42**	
Rel_Vol_ Range	**50**	**43**	
Cell_ Volume_Map	**51**	**44**	
Clustering	**52**	**45**	
Non_ Uniformity	**53**	**46**	
Cell_ Volume	**54**	**47**	
Coarse_ Cell_ Volume	**55**	**48**	**16**
Average_ Cell_ Elongation	**56**	**49**	
Net_Cell_ Elongation	**57**	**50**	
Cell_ Angle_to_Vertical	**58**	**51**	
Cell_Alignment	**59**	**52**	
Vertical_ Elongation	**60**	**53**	
Circulation	**61**	**54**	
Circulation_Horiz_ Offset	**62**	**55**	
Circulation_Vert_Offset	**63**	**56**	
Curvature	**64**	**57**	

Table 8. Spiral White trait data supplied by end users for the 2005 and 2006 harvests

	Spiral White Traits	2006	2005	Reduced 2006
Loaf Volume	Loaf mass (g)	**1**	**1**	
	Loaf volume (ml)	**2**	**2**	**1**
	Specific volume (ml/g)	**3**	**3**	
Crumb colour	L*	**4**	**4**	**2**
	a*	**5**	**5**	
	b*	**6**	**6**	
	X	**7**	**7**	
	Y	**8**	**8**	
	Z	**9**	**9**	
Day 1 TPA	firmness(g)	**10**	**10**	**3**
	adhesiveness	**11**	**11**	
	springiness	**12**	**12**	
	cohesiveness	**13**	**13**	
	gumminess (g)	**14**	**14**	
	chewiness (g)	**15**	**15**	
	resilience	**16**	**16**	
Day 3 TPA	firmness(g)	**17**	**17**	**4**
	adhesiveness	**18**	**18**	
	springiness	**19**	**19**	
	cohesiveness	**20**	**20**	
	gumminess (g)	**21**	**21**	
	chewiness (g)	**22**	**22**	
	resilience	**23**	**23**	
C-Cell	Slice Area / px	**24**	**31**	
	Height (max) / px	**25**	**32**	**5**
	Height (avg) / px	**26**	**33**	
	Breadth / px	**27**	**34**	
	Height / Breadth	**28**	**35**	
	Wrapper Length / px	**29**	**36**	
	Total Concavity / %	**30**	**37**	
	Left Concavity / %	**31**	**38**	
	Right Concavity / %	**32**	**39**	
	Top Concavity / %	**33**	**40**	
	Bottom Concavity / %	**34**	**41**	
	Left Break / %	**35**	**42**	
	Right Break / %	**36**	**43**	
	Left Break Height / px	**37**	**44**	
	Right Break Height / px	**38**	**45**	
	Left Break Depth / px	**39**	**46**	
	Right Break Depth / px	**40**	**47**	
	Left Break Position / px	**41**	**48**	
	Right Break Position / px	**42**	**49**	
	Top Left Shoulder	**43**	**50**	

Table 8. (contd) Spiral White trait data supplied by end users for the 2005 and 2006 harvests

	Spiral White Traits	2006	2005	Reduced 2006
C-Cell	Top Right Shoulder	**44**	**51**	
	Bottom Left Roundness	**45**	**52**	
	Bottom Right Roundness	**46**	**53**	
	Slice Brightness	**47**	**54**	
	Cell Contrast	**48**	**55**	
	Number of Cells	**49**	**56**	**6**
	Number of Holes	**50**	**57**	
	Area of Cells / %	**51**	**58**	
	Area of Holes / %	**52**	**59**	
	Volume of Holes	**53**	**60**	
	Wall Thickness / px	**54**	**61**	**7**
	Cell Diameter / px	**55**	**62**	**8**
	Cell Vol Range (map)	**56**	**63**	
	Relative Vol Range (map)	**57**	**64**	
	Cell Volume (map)	**58**	**65**	
	Coarse / Fine Clustering	**59**	**66**	
	Non-Uniformity	**60**	**67**	
	Cell Volume	**61**	**68**	
	Coarse Cell Volume	**62**	**69**	**9**
	Average Cell Elongation	**63**	**70**	
	Net Cell Elongation	**64**	**71**	
	Cell Angle to Vertical / °	**65**	**72**	
	Cell Alignment	**66**	**73**	
	Vertical Elongation	**67**	**74**	
	Degree of Circulation	**68**	**75**	
	Circulation Horiz Offset / %	**69**	**76**	
	Circulation Vert Offset / %	**70**	**77**	
	Curvature	**71**	**78**	
Calculated parameters	Firmness 3/LV	**72**		**10**
	Firm.3-Firm.1	**73**		**11**
	Firm.3-Firm.1/LV	**74**		**12**
	left+right concavity	**75**		**13**
	L+R conc./LV	**76**		**14**
	L+R conc./max height	**77**		**15**
	No cells/slice area	**78**		**16**

Table 9. Wholemeal trait data supplied by end users for the 2005 and 2006 harvests

Wholemeal Traits	2006	2005	Reduced set 2006
Dough Consistency Ex Mixer (BU)	**1**	**1**	**1**
Dough Consistency Ex 10 Mins (BU)	**2**	**2**	**2**
Degree of Softening	**3**	**3**	**3**
Mix Time (s)	**4**		**4**
Temp Ex mixer (°C)	**5**		**5**
Temp Ex 10 minutes (°C)	**6**		**6**
Gross Water to 410BU	**7**	**4**	**7**
Manual Stickiness	**8**	**5**	**8**
Proof Height (mm)	**9**	**6**	
Oven Spring (mm)	**10**	**7**	
Loaf Height mm	**11**	**8**	
Loaf volume, ml	**12**	**9**	**9**
L*	**13**	**10**	**10**
a*	**14**	**11**	
b*	**15**	**12**	
Firmness (Day 1), N	**16**	**13**	**11**
Firmness (Day 3), N	**17**	**14**	**12**
Day 3 - Day 1	**18**		**13**
(Day 3 - Day 1) / Volume	**19**		**14**
Day 3 / Volume	**20**		**15**
Resilience (Day 1), %	**21**	**16**	
Resilience (Day 3), %	**22**	**17**	
Slice_Area	**23**	**19**	
Average Height_Max	**24**	**20**	**16**
Height_Avg	**25**	**21**	
Breadth	**26**	**22**	
Height_ Breadth	**27**	**23**	
Wrapper_Length	**28**	**24**	
Total_ Concavity	**29**	**25**	
Left_ Concavity	**30**	**26**	
Right_ Concavity	**31**	**27**	
Left + Right Concavity	**32**		**17**
(Left + Right Concavity)/Loaf Volume	**33**		**18**
(Left + Right Concavity)/Slice Height	**34**		**19**
Top_ Concavity	**35**	**28**	
Bottom_ Concavity	**36**	**29**	
Left_Bk_Area	**37**	**30**	
Right_Bk_Area	**38**	**31**	
Left_Bk_Height	**39**	**32**	
Right_Bk_Height	**40**	**33**	
Left_Bk_Depth	**41**	**34**	
Right_Bk_Depth	**42**	**35**	
Left_Bk_Pos	**43**	**36**	

Table 9. (contd) Wholemeal trait data supplied by end users for the 2005 and 2006 harvests

Wholemeal Traits	2006	2005	Reduced set 2006
Right_Bk_Pos	**44**	**37**	
Top_Left_Shoulder	**45**	**38**	
Top_Right_Shoulder	**46**	**39**	
Bottom_ Left_Roundness	**47**	**40**	
Bottom_ Right_Roundness	**48**	**41**	
Slice_ Brightness	**49**	**42**	
Cell_ Contrast	**50**	**43**	
Number_ of_Cells	**51**	**44**	**20**
Number Cells / Slice Area	**52**		**21**
Number_ of_Holes	**53**	**45**	
Area_of_ Cells	**54**	**46**	
Area_of_ Holes	**55**	**47**	
Volume_ of_Holes	**56**	**48**	
Wall_ Thickness	**57**	**49**	**22**
Cell_ Diameter	**58**	**50**	**23**
Cell_Vol_Range	**59**	**51**	
Rel_Vol_ Range	**60**	**52**	
Cell_ Volume_Map	**61**	**53**	
Clustering	**62**	**54**	
Non_ Uniformity	**63**	**55**	
Cell_ Volume	**64**	**56**	
Coarse_ Cell_ Volume	**65**	**57**	**24**
Average_ Cell_ Elongation	**66**	**58**	
Net_Cell_ Elongation	**67**	**59**	
Cell_ Angle_to_Vertical	**68**	**60**	
Cell_Alignment	**69**	**61**	
Vertical_ Elongation	**70**	**62**	
Circulation	**71**	**63**	
Circulation_Horiz_ Offset	**72**	**64**	
Circulation_Vert_Offset	**73**	**65**	
Curvature	**74**	**66**	

2.7.3 Initial data analysis

Initial statistical analysis of the data was carried out using Minitab 13.1 software. For each year and product, correlation matrices were calculated between traits within products and years, and between major bread making traits across years. As an example, Table 10 summarises the correlation for major traits associated with breadmaking performance across the years for each product and cross.

Table 10. Correlations between major baking parameters across years: Hereward × Malacca data

H×M Spiral White 2005-06			H×M Wholemeal 2005-06			H×M CBP 2005-06		
	Cor.	**Sig.**		**Cor.**	**Sig.**		**Cor.**	**Sig.**
Loaf volume (ml)	0.438	***	Loaf volume, ml	0.343	***	Loaf volume, ml	0.296	**
L*	0.458	***	L*	0.555	***	L*	0.301	**
Firmness (Day 1)	0.496	***	Firmness (Day 1), N	0.410	***	Firmness @25% compression (Day 1), N	0.148	NS
Firmness (Day 3)	0.446	***	Firmness (Day 3), N	0.349	***	Firmness @25% compression (Day 3), N	0.265	*
Height (max) / px	0.39	***	Average Height_Max	0.224	*	Average Height_Max	0.133	NS
Number of Cells	0.448	***	Number_ of_Cells	-0.238	*	Number_ of_Cells	0.145	NS
Wall Thickness / px	0.359	***	Wall_ Thickness	-0.405	***	Wall_ Thickness	-0.046	NS
Cell Diameter / px	0.347	***	Cell_ Diameter	-0.343	***	Cell_ Diameter	-0.099	NS
Coarse Cell Volume	0.442	***	Coarse_ Cell_ Volume	-0.246	*	Coarse_ Cell_ Volume	-0.094	NS
			Dough Consistency Ex Mixer (BU)	0.134	NS			
			Dough Consistency Ex 10 Mins (BU)	0.214	*			
			Degree of Softening	0.108	NS			
			Mix Time (s)	0.211	NS			
			Temp Ex mixer (°C)	0.017	NS			
			Temp Ex 10 minutes (°C)	0.268	*			
			Gross Water to 410BU	0.476	***			
			Manual Stickiness	0.268	*			

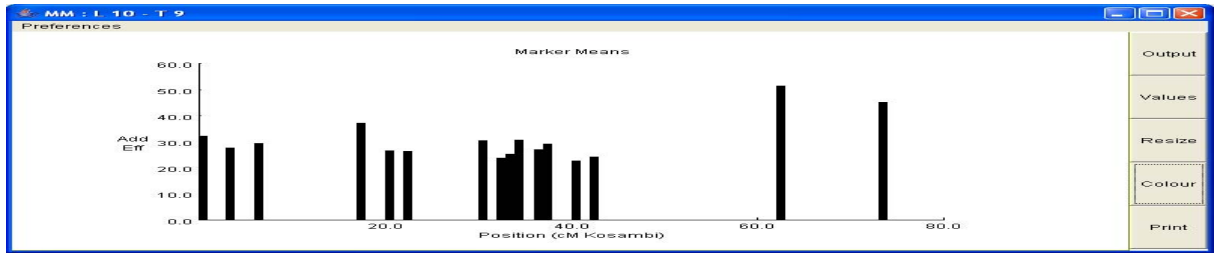
Table 11. Part of the correlation matrix for wholemeal traits for the Malacca x Charger cross in 2005 [only significant correlations (p<0.05) are shown]

	M x C Wholemeal	Dough Consistency Ex Mixer (BU's)	Dough Consistency Ex 10 Mins (BU's)	Degree of Softening	Mix Time (s)	Temp Ex mixer (°C)	Temp Ex 10 minutes (°C)	Gross Water to 410BU's	Manual Stickiness	Proof Height (mm)	Oven Spring (mm)	Loaf Height mm	Loaf volume, ml
		1	2	3	4	5	6	7	8	9	10	11	12
2	Dough Consistency Ex 10 Mins (BU's)	0.849											
3	Degree of Softening		-0.563										
4	Mix Time (s)	-0.478	-0.418										
5	Temp Ex mixer (°C)												
6	Temp Ex 10 minutes (°C)		-0.314	0.351		0.369							
7	Gross Water to 410BU's				0.481	-0.341	-0.388						
8	Manual Stickiness	-0.304	-0.263		0.215		-0.22						
9	Proof Height (mm)												
10	Oven Spring (mm)		-0.277	0.252						-0.402			
11	Loaf Height mm	-0.28	-0.401	0.32	0.218			-0.211		0.268	0.775		
12	Loaf volume, ml				0.254					0.429	0.324	0.637	
13	L*							-0.36			0.383	0.299	0.571
14	a*							0.292			-0.489	-0.384	-0.575
15	b*										-0.285	-0.233	-0.206
16	Firmness (Day 1), N				-0.266			-0.201		-0.426			-0.264
17	Firmness (Day 3), N				-0.28			-0.254		-0.407	0.217		-0.218
18	Firmness (Day 7), N	0.202			-0.261		0.225	-0.309		-0.447	0.265		
19	Resilience (Day 1), %		-0.26	0.274							0.614	0.538	0.477
20	Resilience (Day 3), %		-0.32	0.302	0.229						0.543	0.524	0.485

Figure 1. An example of output from QTL Café showing the Single ANOVA and Interval Mapping QTL results

Malacca x Charger, 2005: Wholemeal Loaf Volume Linkage Group 10

(lg 10) (Marker Positions and allele differences)



QTL Location by Interval mapping

QTL Mapping by Interval Mapping: Result 1 QTL

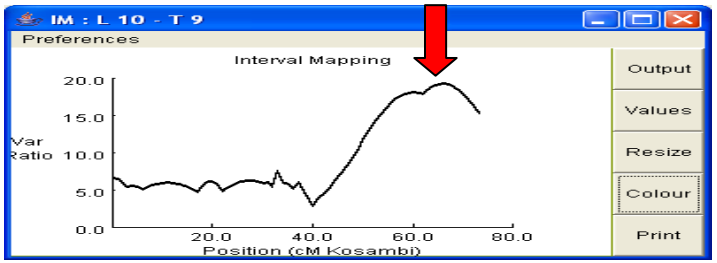
Linkage Group: 10

Trait: 9

QTL located at 66 cM

Test Statistics : F 19.3566 LR 18.0006

Res. SS - Full Model	1154373.6	d.f.	96
Res. SS - Red. Model	1387131.1	d.f.	97
Mean	3103.163		



As an example of correlations between traits within years, Table 11 shows part of the large correlation matrix between traits for wholemeal performance in the Malacca x Charger cross 2005 harvest.

2.7.4 QTL analysis

For QTL discovery, an eclectic approach was used for all data sets, from single marker ANOVA through to interval mapping and marker regression approaches using QTL Café Software (http://www.biosciences.bham.ac.uk/labs/kearsey). Most of the results from the different approaches concurred for QTL location to within a few cM. Probabilities of less than 0.05 were taken as a lower significance level, but at low significance levels only QTL detected over the two seasons were given credence and tabulated. Figure 1 shows an example of the QTL Café output for the trait loaf volume measured in 2005 for wholemeal in the Malacca × Charger cross.

For the data from the 2005 harvest, all traits were analysed for the location of QTL. However, from these results it was apparent that many of the traits were highly correlated, producing co-locating QTL. Thus, to reduce the number of parameters that required QTL analysis in 2006, a sub-set of the most important and interesting (based on the 2005 analysis) only was analysed (see Tables 5-9). However, additionally in 2006 and retrospectively for the 2005 data, new derived traits were also analysed.

QTL data were summarised into Excel spreadsheets for each chromosome and trait within years and across years. The locations of all QTL within a cross, across traits and years, were tabulated onto the genetic maps to produce an overview of all new QTL found, and particularly of co-locating QTL for different traits and different products. Comparative maps between the crosses for the same chromosomes were also prepared using cross-referencing genetic markers.

2.8 Metabolomics

2.8.1 Extraction procedure

3 x 30mg of each flour sample plant material was weighed into an autoclaved 2ml eppendorf tube. 1ml 80:20 D_2O:CD_3OD containing 0.05% w/v d_4 TSP (sodium salt of trimethylsilylpropionic acid) was added to each sample. The contents of the tube were mixed thoroughly and then heated at 50^{o}C in a water bath for 10 minutes followed by

2 minutes at 90°C. Samples were then left for 45 minutes to cool. After cooling, the samples were spun down in a micro-centrifuge for 5 minutes. 750µl of the supernatant was added to a 5mm NMR tube. All spectra were acquired under automation at a temperature of 300K on a Bruker Avance spectrometer operating at 600 MHz ^{1}H observation frequency using a selective inverse probe and the WATERSUP pulse sequence with a relaxation delay of 5 seconds. Each spectrum consisted of 128 scans of 32k data points with a spectral width of 4845Hz. The spectra were automatically Fourier transformed using an exponential window with a line broadening value of 0.5Hz, phased and baseline corrected within the automation software. ^{1}H NMR chemical shifts in the spectra were referenced to d_4-TSP at δ0.00.

2.8.2 Data reduction of the NMR spectra

The ^{1}H NMR spectra were automatically reduced to ASCII files using AMIX (Analysis of MIXtures software v.3.0, Bruker Biospin). Spectra were scaled to d_4 TSP and reduced to integrated regions or "buckets" of equal width (0.01 ppm) corresponding to the region of δ9.995 to δ-0.5. The regions between δ4.855 and δ4.775 were removed prior to statistical analyses, thus eliminating any variability in suppression of the water sample. The signals corresponding to d_4 methanol(δ3.325-δ3.295) and d_4 TSP (δ0.00) were also removed at this stage. The generated ASCII file was imported into Microsoft Excel for the addition of labels and then imported into SIMCA-P 11.0 (Umetrics, Umea, Sweden) for multivariate analysis.

2.8.3 Multivariate analysis

Data were analysed using SIMCA-P 11.0 (Umetrics, Umea, Sweden). All data were mean-centre scaled. Principal components analysis was carried out on all data sets.

2.8.4 Correlation analysis

Data was imported into Spotfire for ANOVA calculations and to examine correlations with supplied QTL data.

2.9 Fourier transform infrared spectroscopy

Spectra of dry flours and flours hydrated with H_2O were obtained on a BioRad FTS 165 FT-IR spectrometer using a Golden Gate single reflection ATR (attenuated total reflectance) sampling system. 256 scans at 2 cm^{-1} resolution were co-added and the

spectra were corrected for water vapour signals and by subtraction of the water signal in the 1650 cm^{-1} region.

Spectra of flours hydrated with D_2O were obtained using a Bruker IFS 66S and the Golden Gate ATR attachment. 64 spectra at 2 cm^{-1} resolution were co added and no further corrections to the spectra were made.

Flours were hydrated with H_2O as follows: to 0.5 g of flour in a sample tube 1 ml of water was added and the sample left for 24 hours. Excess water was removed by pipette and the sample placed on the ATR plate. In the case of D_2O hydration the excess was removed after standing overnight and second aliquot added. This was left to stand for 24 hours. The excess water was removed, a further aliquot added and the spectra run with excess fresh D_2O.

3. Results and discussion

3.1 Development of a wheat quality map

Both flour and baking quality data were subjected to PCA to assess how samples were grouped by both variety (13 tested) and growing location (5 tested). In both cases, the pattern of variation showed that the first few principal components described the majority of the variation in the data for flour and bread quality (Figures 2 and 3 respectively).

Based on these results, the first few principal components in each case were used to assess the distribution of samples in principal component space (Figures 4 and 5). In all cases, clear distinctions between the different groups represented by the samples were not evident. However, the way in which samples were clustered by growing location appears more tangible than that for varieties, particularly for scores 1 and 2. For the flour parameters, Soissons was slightly removed from the other samples, having low values of score 1 allied to higher values of score 2. This direction was related to higher values of gel protein mass and G', which would be consistent with the stronger gluten properties exhibited by Soissons. The clustering by variety was even less apparent when assessing baking performance with individual samples for each variety at different sites being distributed throughout the principal component space defined by the entire population.

Figure 2. Cumulative variation described by principal components for flour properties

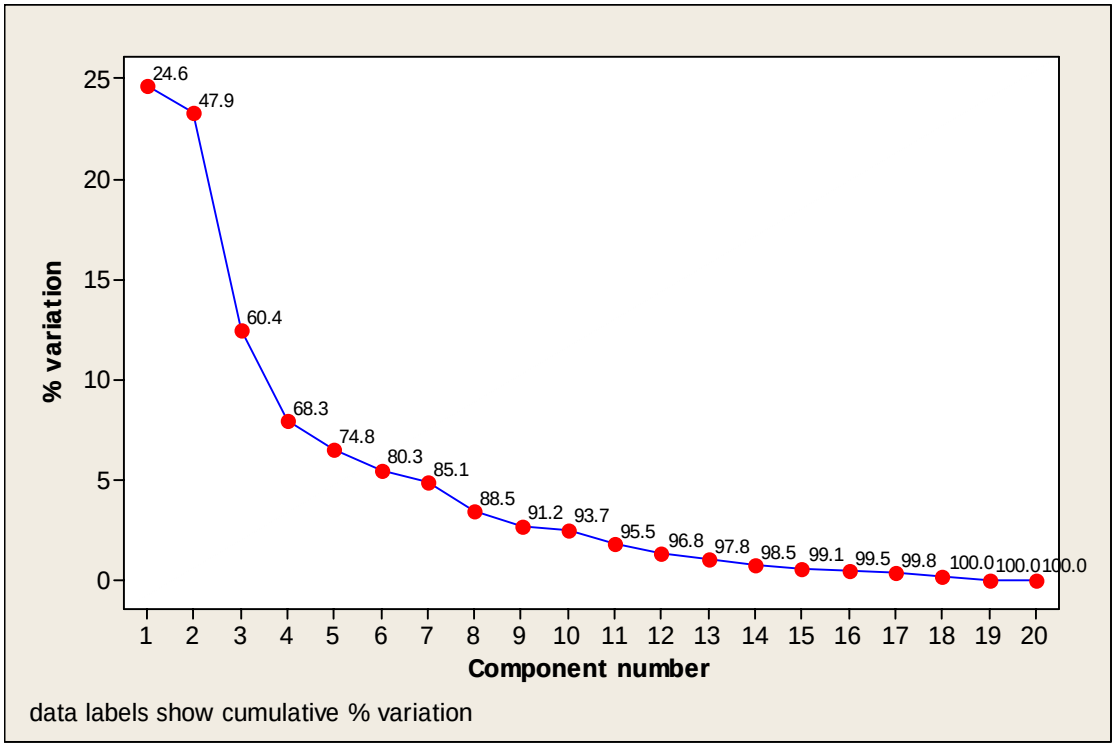


Figure 3. Cumulative variation described by principal components for baking performance

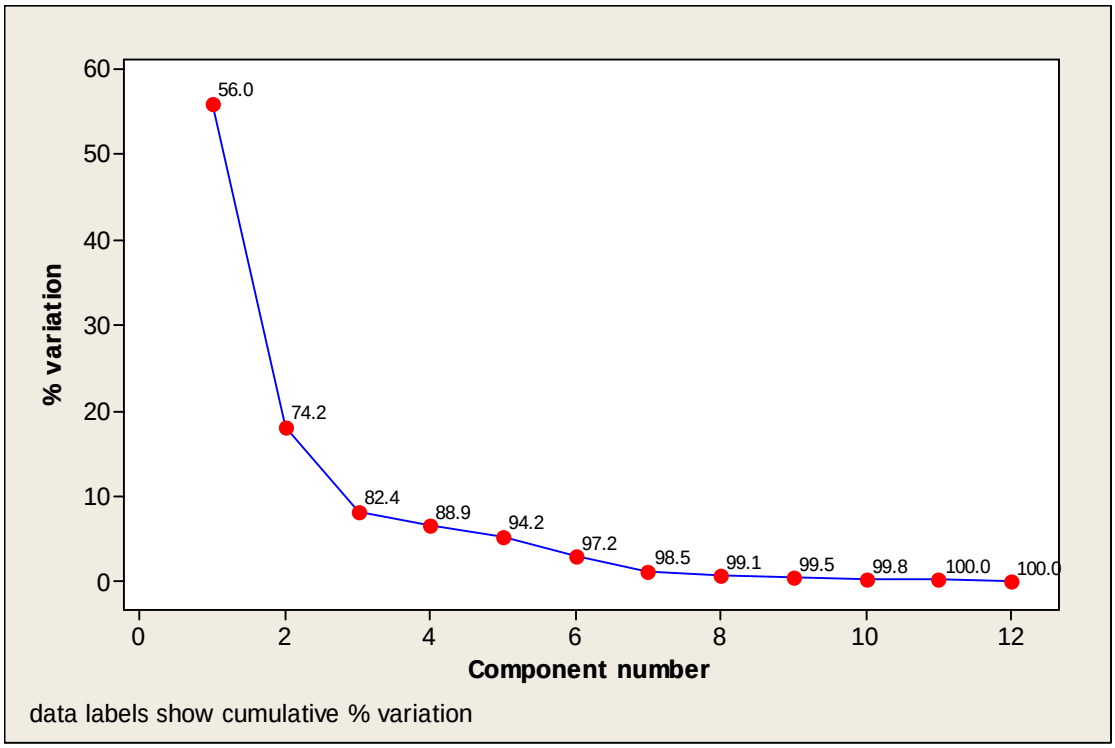
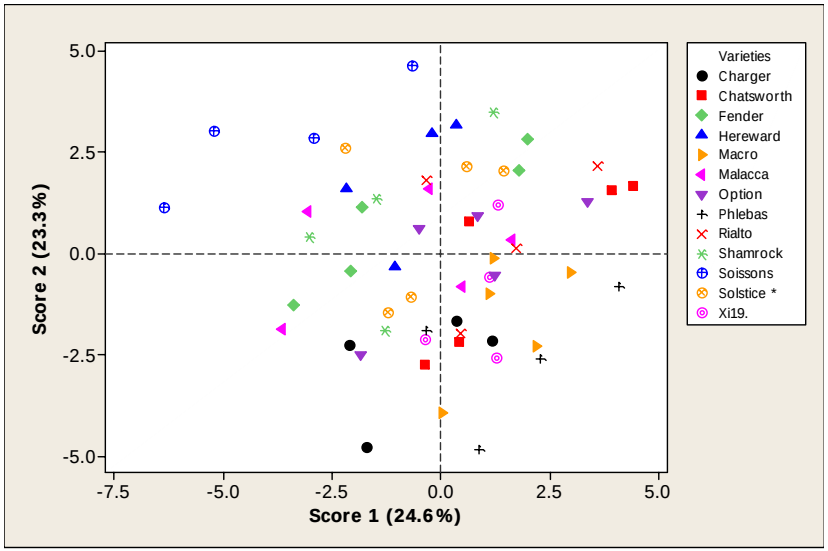
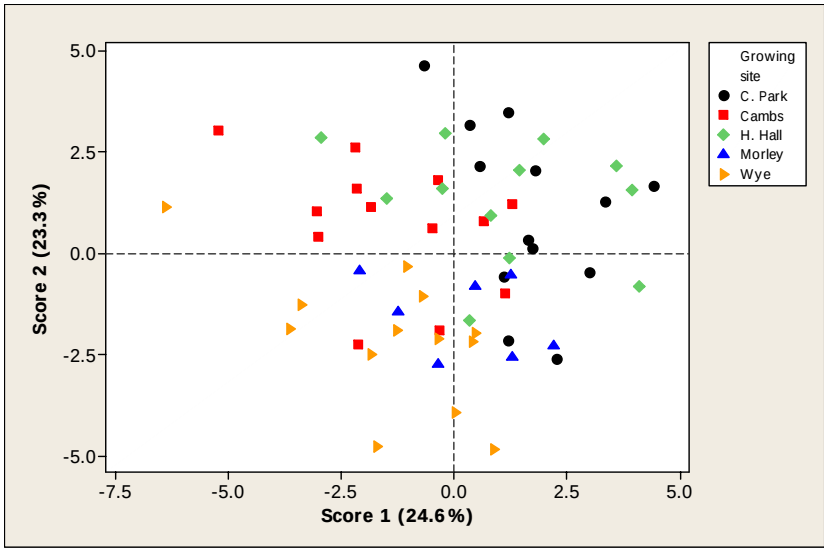
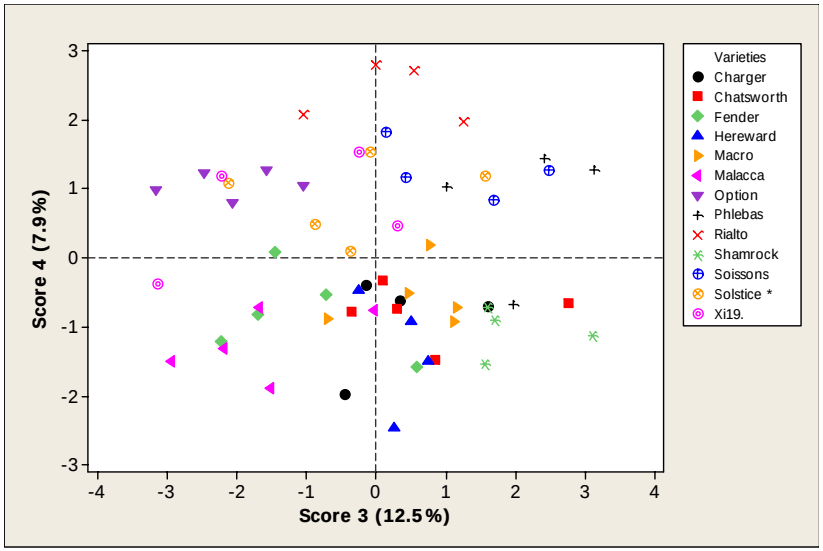


Figure 4. Principal components analysis based on flour properties







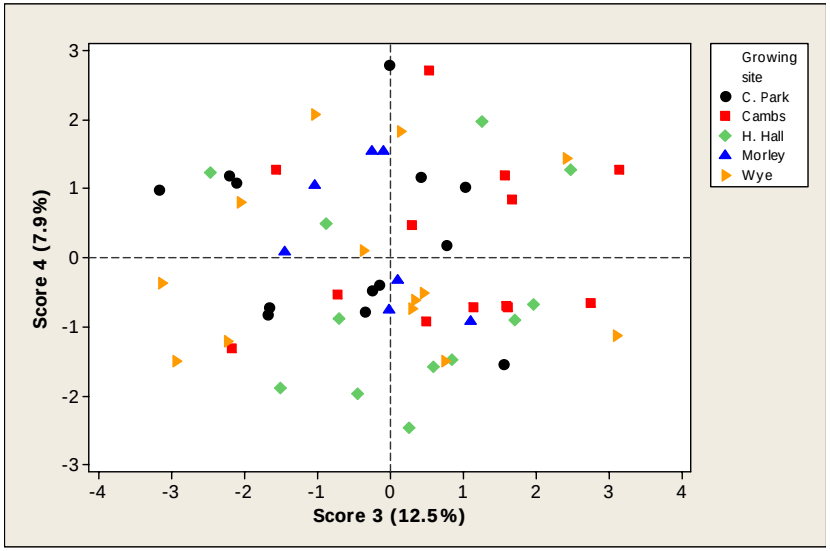
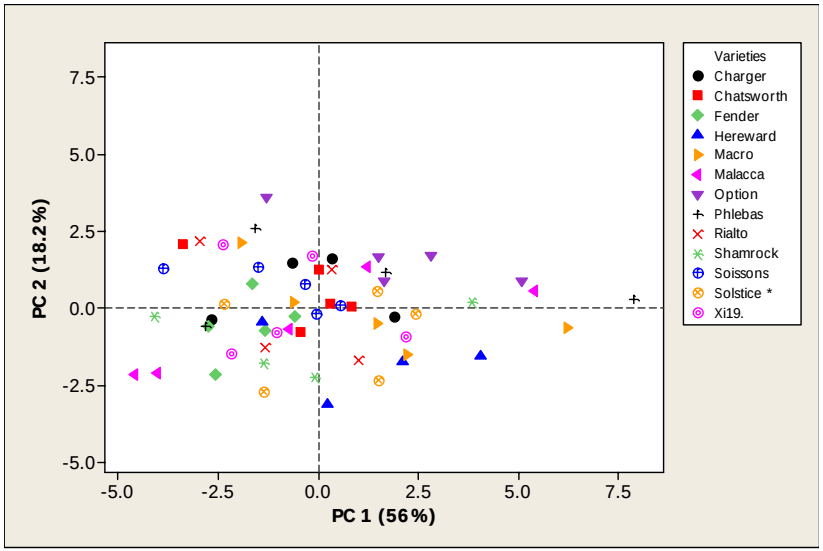
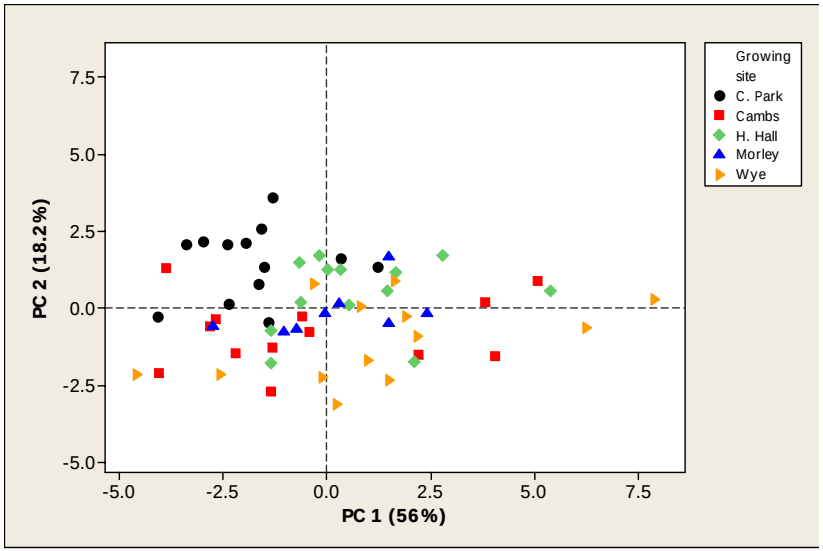
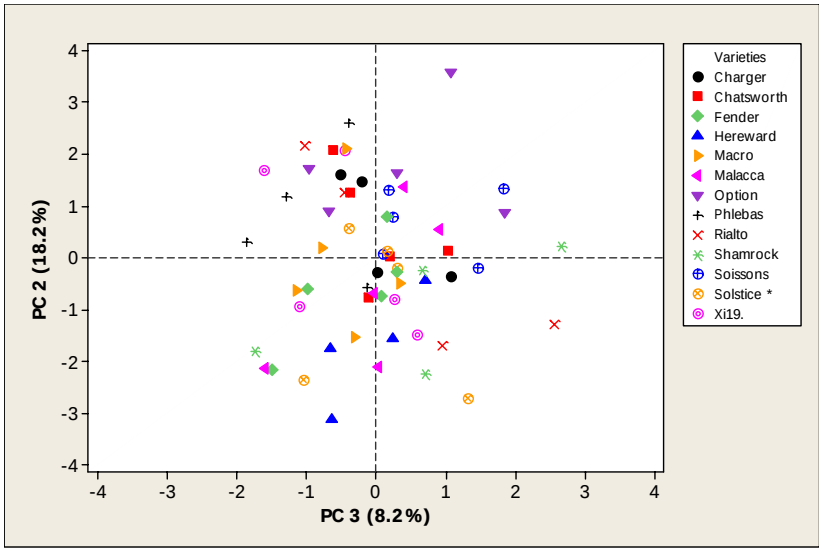
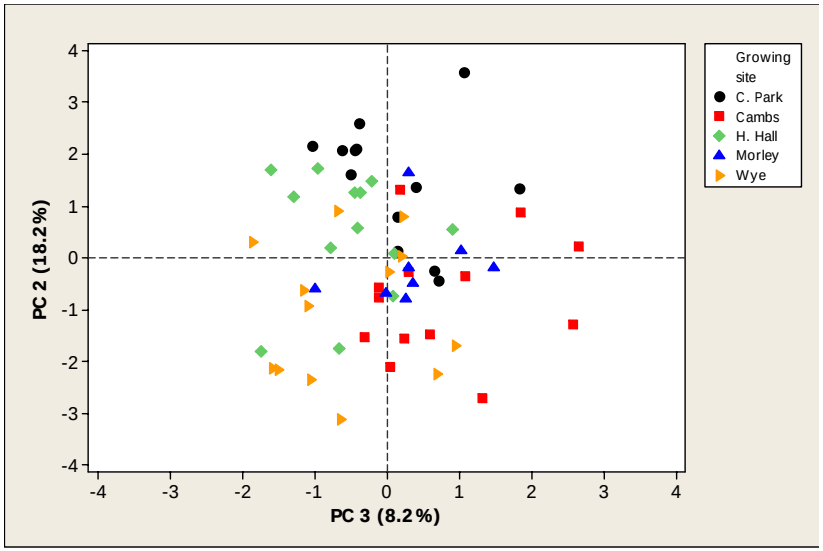


Figure 5. Principal components analysis based on baking performance









The different sites showed some progressive differences in flour properties when assessing principal components 1 and 2. With the exception of the Cambridge (Cambs) site, the different growing locations were broadly distributed across an axis from the upper right quandrant to the lower left. This direction was associated with parameters related to water absorption with those further to the upper right of the graph being higher in this and associated parameters. The positioning of the samples from the Cambridge site indicated that the samples grown there tended to have stronger gluten characteristics as assessed using gel protein analysis.

For bread properties, the different sites were less widely distributed across the principal component space and the differences between them were less marked. However, samples from the Cockle Park (C. Park) site tended to reside in the upper left quandrant which indicated higher cell numbers but lower loaf volume when compared with the other growing locations.

In addition to understanding how the different samples clustered in relation to defined quality attributes, the assessments carried out on both wheat and flour were used to predict the properties of the final baked loaves generated from them. The number of factors used in predicting baking parameters from grain data was 6, which represented 96.4% of the X variance. When using flour parameters, 7 components were used, representing 79.8% of the X variance. The results of the calibration and predictive performance in terms of squared correlation coefficients are given in Table 12.

The results confirm the complexity of baking performance of wheat flour with the wheat and flour quality tests used only giving some indication of performance rather than being accurate predictors. This agrees with previous work (Millar, 2003) and is not only due to the complexity of the biochemical and physical interactions underpinning breadmaking but also relates to the inherent difficulty in accounting for interactions within intermediate processes which also contribute to baking performance.

Table 12. Prediction of individual loaf attributes from grain and flour assessment

	Grain characteristics		Flour characteristics	
Loaf attribute	**R^2 (fit)**	**R^2 (pred)**	**R^2 (fit)**	**R^2 (pred)**
CBP loaf volume	39.0	19.5	59.7	28.3
CBP loaf score	26.2	2.9	35.4	0.0
Number of cells	25.2	4.4	26.1	0.0
Cell area fraction	18.8	0.0	32.1	0.0
Cell wall thickness	31.2	11.1	25.5	0.0
Cell contrast	40.7	22.4	45.7	10.6
Characteristic cell area	21.3	0.0	35.0	0.0
Large cell area	25.2	4.0	42.1	0.0
Cell area variability	25.8	4.8	39.3	0.0
Weighted cell area	27.9	7.0	33.1	0.0
Average cell area	23.8	2.6	26.7	0.0
Spiral loaf volume	31.7	5.9	60.0	41.8

Table 13. Prediction performance using grain parameters

Coefficients	CBP loaf volume	Cell contrast
Specific weight	-0.384	-0.139
Grain length	0.037	0.050
Grain breadth	-0.189	-0.109
Grain perimeter	-0.016	0.028
Grain area	-0.044	-0.005
SKCS grain mass	0.047	0.075
SKCS grain diameter	0.109	-0.112
SKCS hardness	-0.352	-0.469
Grain protein as is NIR	0.103	0.024
Grain protein db NIR	0.159	0.070
Grain moisture NIR	0.401	0.368

When taking account of both sets of equations, it is clear that the most interesting relationships are those used in the prediction of CBP loaf volume, cell contrast (for grain parameters) and no time dough (Spiral) loaf volume (for flour parameters). The standardised weightings of each of the underlying contributors to these equations are given in Tables 13 and 14 where larger absolute values indicate importance in the prediction.

Table 14. Prediction performance using flour parameters

Coefficients	CBP loaf volume	Spiral loaf volume
Milling yield	0.116	0.348
Flour particle size distribution	-0.205	-0.280
Flour moisture (NIR)	0.041	0.013
Flour protein as is NIR	0.176	0.155
Flour protein 14% mb NIR	0.177	0.128
Flour damaged starch	-0.355	-0.193
Flour Hagberg Falling Number	-0.003	0.053
Flour Grade Colour	-0.211	-0.244
Branscan bran	-0.006	-0.077
Branscan specks	0.210	0.204
Farinograph Water Absorption	-0.191	-0.027
Farinograph Development Time	-0.082	-0.107
Farinograph Stability	0.052	0.092
Farinograph Degree of Softening	0.071	0.426
Total pentosans	0.083	0.145
Soluble pentosans	0.386	0.287
Insoluble pentosans	-0.021	0.065
Ceralpha	0.028	0.092
Gel protein weight	0.125	0.188
Gel protein G'	-0.113	-0.406

The major contributors to the predictions of baking performance from grain parameters were grain specific weight, grain hardness and grain moisture. The first two of these may be thought of as parameters which will have an effect on flour properties as a function of milling performance and so there is some rationale for their

inclusion. It is difficult, however, to understand why grain moisture content (as assessed by NIR) should have an impact on subsequent baking performance.

For the flour properties, the loadings were different in relative values for the different test baking methods. For CBP loaf volume, important predictors were soluble pentosans, damaged starch, Branscan specks and flour colour. For no time dough (Spiral) bread, Farinograph Degree of Softening, gel protein G', milling yield and soluble pentosans were the terms having the highest importance in the equation. While all of these parameters would be expected to have some effect, it is clear that the overall levels of prediction are relatively low and that there is no evidence of one particular measurement which is critical in understanding likely baking performance. This underlines once again the difficulties inherent in understanding the factors affecting baking performance.

Given the limited success in separating varietal characteristics on the basis of the standard wheat and flour assessment techniques allied to the poor predictive performance when using these parameters to determine likely baking properties, it is clear that the rationale behind the project in attempting to determine the underlying genetic basis for individual baking parameters is sound. While routine wheat and flour quality parameters are valuable in generating a picture of likely suitability of different flours for particular end uses, the values they give may not be used effectively to predict functionality alone. The approach of generating direct links between wheat genetics and objective measures of baked product quality is thus likely to be more powerful than any potential predictive techniques.

The work has also shown the impact of growing location from which likely differences in performance with differing harvest years may be inferred. It is important, therefore, that any QTL generated should be robust across such environmental effects to ensure that they will be more broadly applicable. Nevertheless, these results tend to indicate that this element of the project will be challenging and that even sound genetic differences will still be affected to a significant degree by growing location and probably harvest year.

3.2 Development of standard methods of bakery product assessment

3.2.1 Bread ring trial

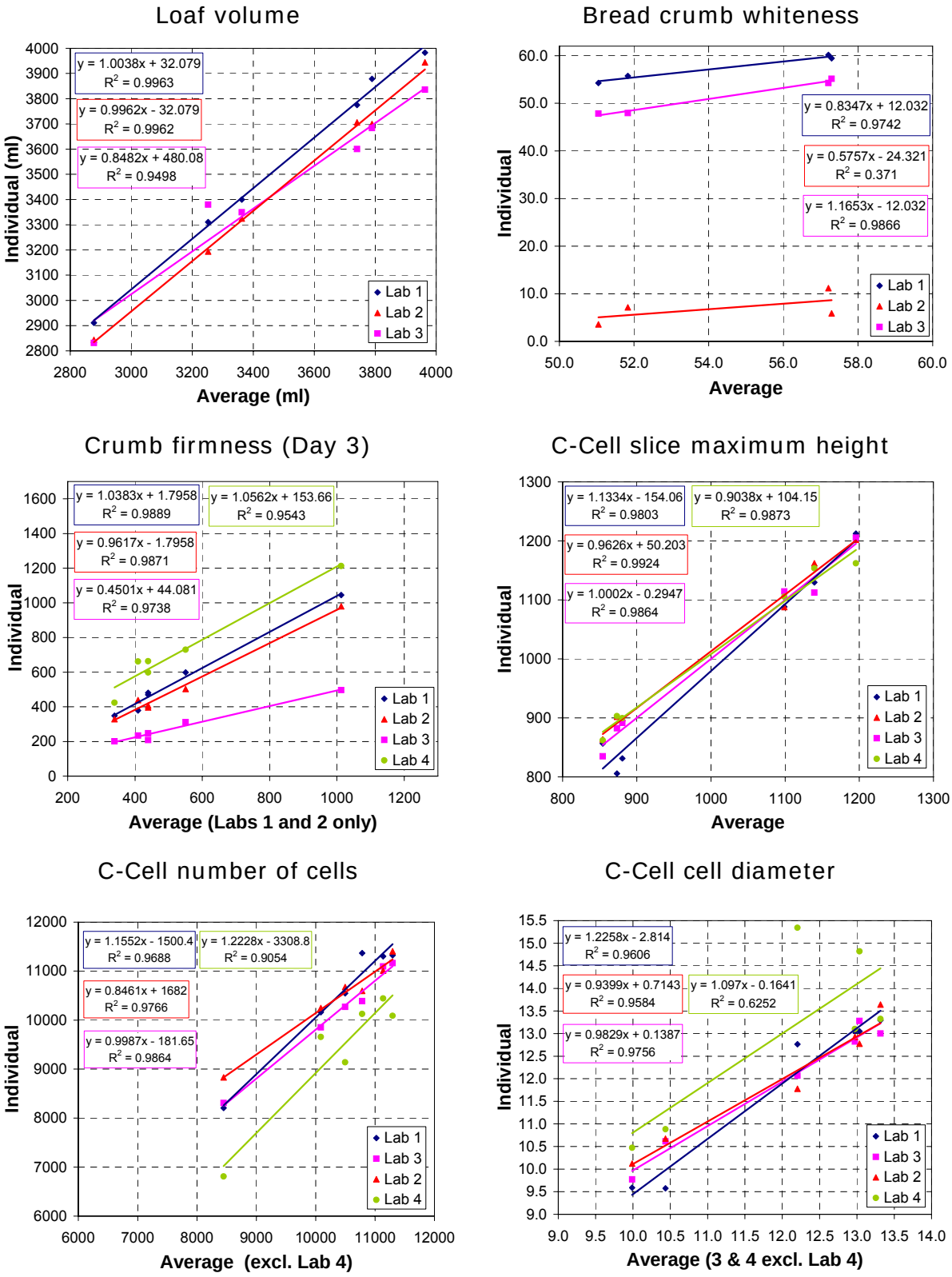
Results for each of the main groups of parameters were used to assess the agreement between the laboratories; there are examples of each in Figure 6. Each graph represents one individual laboratory compared with the mean for all participating laboratories. Not all laboratories submitted assessment of all parameters but the numbering of the laboratories is consistent throughout.

For loaf volume, Laboratory 1 exhibited slightly higher values than the other two laboratories involved and subsequent internal investigation showed why this was so and resolved the issue in advance of further project work. While the mean results obtained in Laboratories 2 and 3 were very close, it was clear that Laboratory 3 had a mild skew in its response. This was also investigated and was shown to be a function of the calibration method then in place. Changes to this were also made to ensure more consistent response across the range of measurement.

Laboratories 1 and 3 both used Tristimulus Y as the response indicating 'whiteness' in the measured bread crumb (white bread only) as this had been requested in the original method. Laboratory 2, however, used an internal method for bread analysis which displayed whiteness results on a different scale. It was agreed that all groups would take a consistent approach for the subsequent end user assessment and report the results obtained for X, Y, Z, L*, a* and b*. While the overall trends shown in each laboratory were similar, it was also recognised that there had been some differences at each site in how the bread had been sliced and presented to the colorimeters. As a result, a global method for the preparation for samples for crumb colour measurement was agreed and used by each participant.

The results for crumb firmness from TPA showed some differences between laboratories with the values for Laboratory 3 showing a skewed response compared with the others. There was also some evidence of slight bias differences between the other end users. Given that each end user would be assessing one product only, small bias differences between groups was not likely to be problematic but further investigation was carried out for the laboratory with skewed results. It was shown that the reason for these results was the use of a constant compressed height with a different thickness of starting material.

Figure 6. Representative parameters determined for bread assessed at end user laboratories



As a result the degree of compression differed causing the difference in the slope of the instrument response to increasing firmness. As a consequence, it was recognised

that there were a number of practical considerations during crumb firmness assessment which needed to be taken into account in the global method to be adopted for bread analysis.

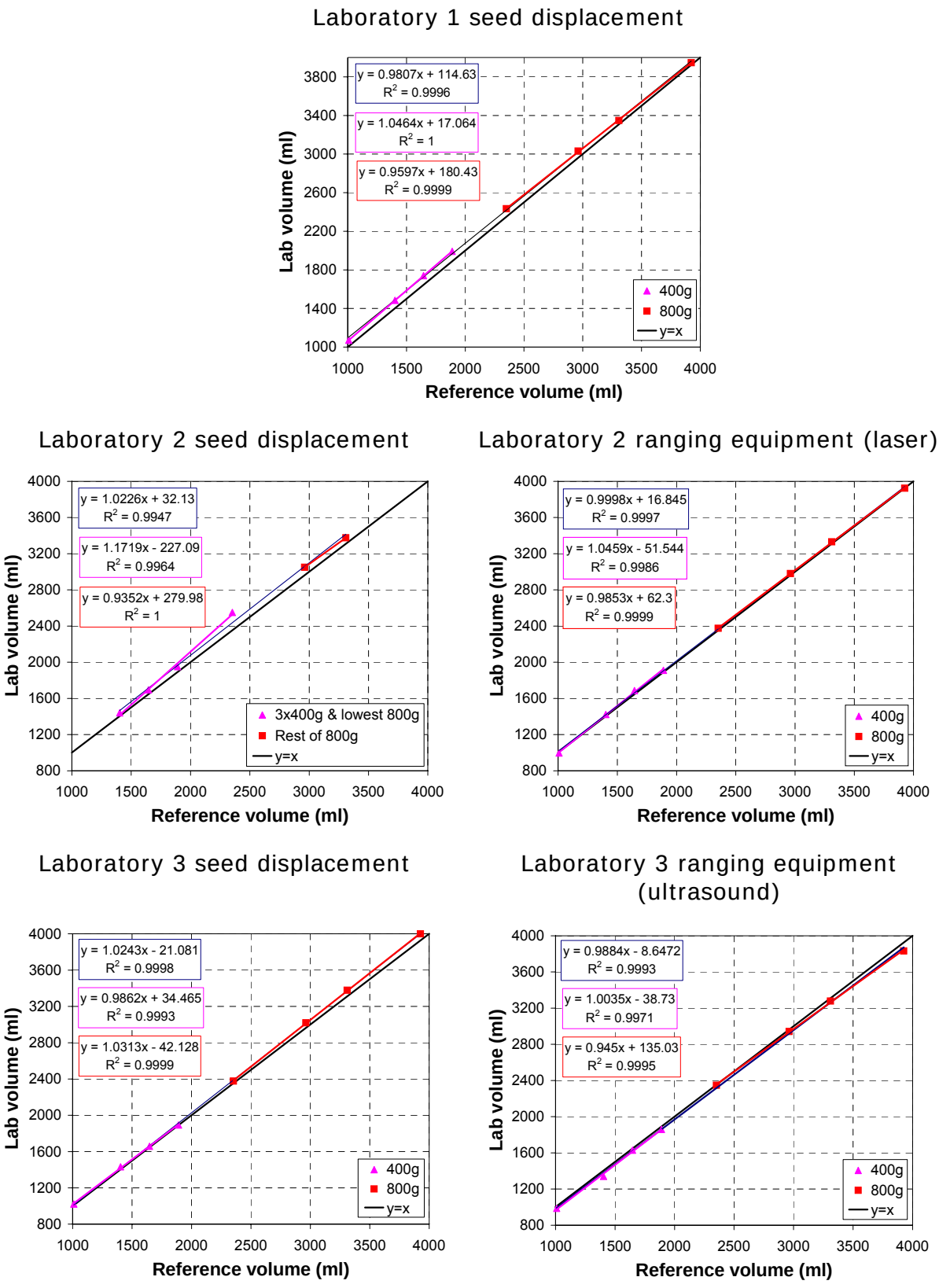
Representative C-Cell parameters showed that there were some minor differences between users for slice dimensions and a more noticeable difference for one laboratory in assessing cell parameters. In the former case, Laboratory 1 had experienced some problems with positioning of slices depending on slice size and the differences in orientation were believed to have caused some of the slight differences observed. Laboratory 4 showed some consistent differences in cell number and size. The results for slice maximum height were consistent with the other groups, however, and so this seemed to represent some issues with cell measurement rather than an overall problem with the way in which the instrument was measuring the slices presented. It was considered that such differences could have resulted from non-flat surfaces on the bread slice and so further evaluation of the slicing method and the frequency of sharpening of the slicer blades was undertaken to resolve this issue.

Overall, the ring trial was successful in highlighting where problems exist for particular methods as well as in generating some basic confidence in the underlying consistency of response between sites. Where differences were observed, subsequent investigation had generally indicated the underlying cause and allowed for remedial action before generating results for the doubled haploid lines.

3.2.2 Loaf volume calibration

The results (Figure 7) extend the previous work on bread, taking into account changes to internal procedures as a result of the initial ring trial. It is clear that the results obtained demonstrate acceptable consistency in response between the different laboratories as well as for the different approaches to measuring loaf volume which were available (seed displacement and ranging by ultrasound or laser). The configuration of the ranging equipment at Laboratory 1 was such that the reference loaves could not be used without significant changes to their volume for subsequent determinations. Nevertheless this method had been internally validated by that group to allow it to be used for the main part of the project.

Figure 7. Performance of ranging and seed displacement loaf volume techniques assessed using standard reference loaves (regression statistics in upper box are for entire dataset)



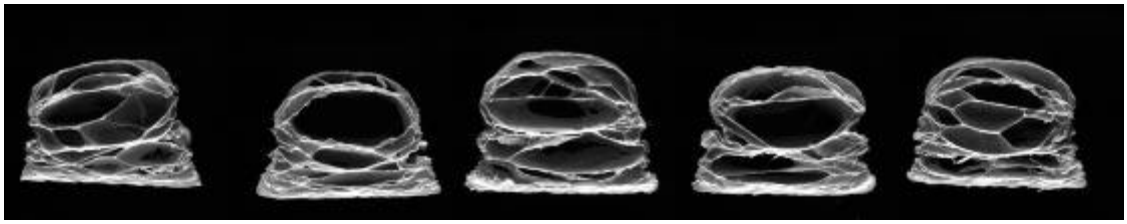
Subsequent to the first year's assessment, further changes were made to the system for loaf volume determination for this laboratory and this was validated using the same reference loaves prior to undertaking the second year's analysis.

For the laboratories where both methods of determining loaf volume were assessed, the ranging equipment exhibited very good performance. This was particularly so for the laser-based system which represents an upgraded version of the original ultrasound-based device. There had been some initial concerns in developing suitable bread mounting systems for use with the TexVol ranging systems as it was postulated that the particularly soft characteristics of UK tinned bread would allow a degree of 'sag' when suspended by the system for measurement. In turn, this was considered to give potential for erroneous values to be recorded. The reference loaves were stiff and so were not prone to such problems. Nevertheless, once an appropriate method for mounting real loaves was developed, Laboratory 2 moved to the ranging system for all the subsequent product assessment within the project. Laboratory 1 also used the TexVol system for 2006 harvest samples. For Laboratory 3, the results for seed displacement and ranging showed broadly similar levels of performance with slight but opposing levels of skew, particularly for the 800g sized reference loaves. On this basis, it was felt more practical to continue with seed displacement within the project given the familiarity with this approach for the various operators.

3.2.3 Puff pastry method development

The production method minimised as far as possible the inherent variability within puff pastry production as illustrated by Figure 8.

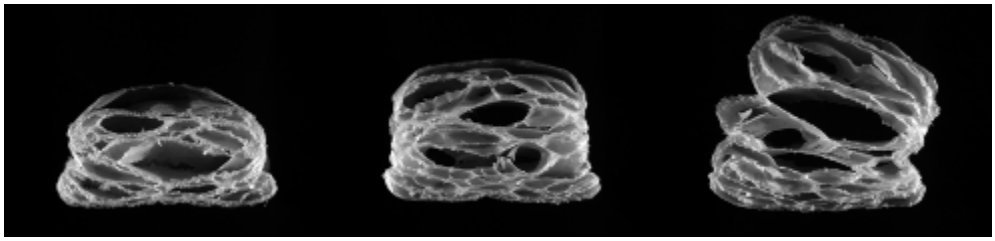
Figure 8. Replicate samples of puff pastry from a single flour sample



The method development exercise resulted in the use of five replicates for each sample with the average of these being quoted in each case.

The ability of the production method to discriminate between samples expected to produce different qualities of puff pastry may also be seen in Figure 9. In this case, flour protein content and type have been used as the main factors affecting puff pastry quality and the results show good discrimination.

Figure 9. Differences in puff pastry quality as a function of flour type



10% protein biscuit flour 11% protein breadmaking flour 13% protein breadmaking flour

Given the range of parameters which impact on acceptability or otherwise of puff pastry, it was recognised that in addition to the C-Cell data, an overall, objective, quality score was required. This was achieved through the use of 3 sets of pastry samples. The first set (52) was assessed by C-Cell and 2 expert assessors. In addition to scores, the expert assessors also detailed their reasoning behind the scores. The scores were used to generate a model at CCFRA within which terms related to fineness of structure, layering, cell size and interactions between fineness and layering were selected as good predictors of the expert assessors' scores. This model was then used to calculate an objective scoring system using a weighted formula containing direct and derived C-Cell outputs. The validation exercise showed that while the objective scores were basically reliable, there were differences from the results generated by expert assessors in 2 circumstances:

- When a lack of symmetry was shown.
- When the structure was dominated by a single large bubble.

Modifications were then made to the weighted formula to include these parameters and a further validation process undertaken. This provided evidence that the reliability of the modified calculated score was at least as good as that shown between experts. On this basis, the method was used for all of the puff pastry assessment undertaken throughout the project.

3.3 Wheat and flour quality

Basic wheat quality parameters were shown to vary within populations, between populations and between harvest years. Table 15 summarises the results for a selection of parameters; the complete set of data is included in the electronic output of the project circulated to members of the project consortium.

Table 15. Basic wheat quality parameters from the sample set used in the project for QTL analysis

	Protein content (%) dmb		Grain hardness index by NIR		Hagberg Falling Number (s)	
	Mean	Range	Mean	Range	Mean	Range
	2005 harvest					
H×M	14.4	13.5-15.2	72	56-83	322	169-418
M×C	13.1	11.8-14.8	60	45-70	351	236-440
S×S	12.8	12.1-13.8	70	44-93	211	98-302
	2006 harvest					
H×M	14.6	13.5-15.7	65	57-74	405	352-443
M×C	13.4	12.2-14.4	65	47-81	417	288-492
S×S	15.1	13.9-16.9	77	61-93	385	312-450

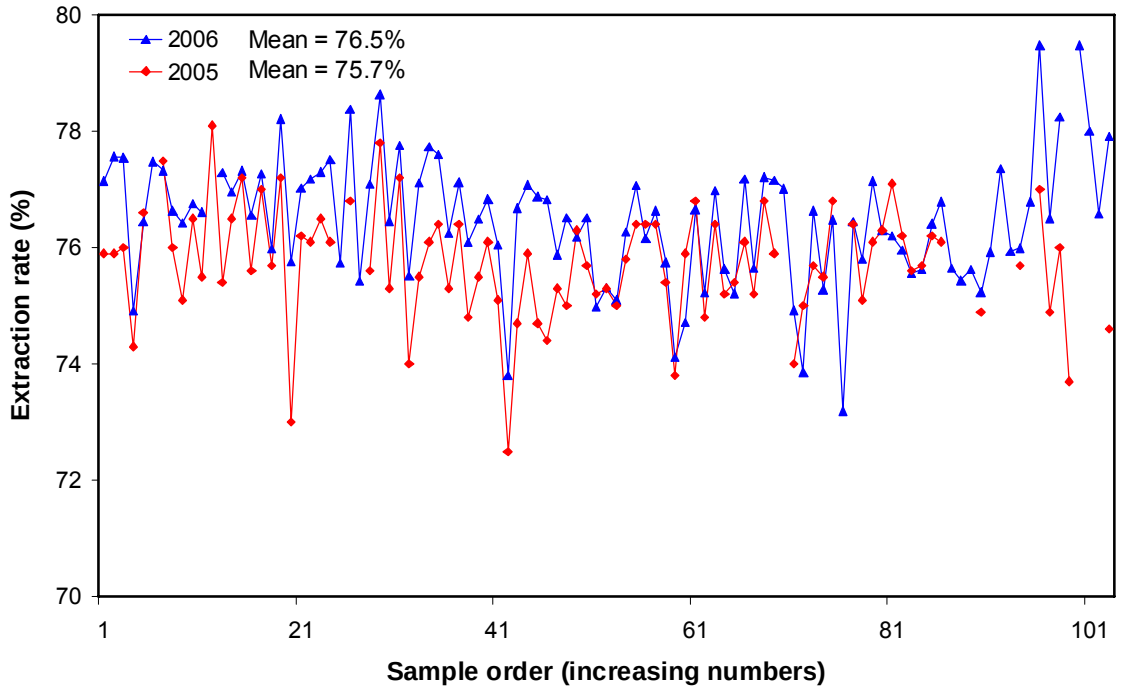
H×M = Hereward × Malacca population, M×C = Malacca × Charger population, S×S = Shango × Shamrock population

The different growing seasons had a significant effect on grain quality with values for Hagberg Falling Number being lower in 2005. This was primarily due to 2006 having an unusually dry early period of harvest within which all samples were combined. The Shango×Shamrock population also showed strong variation between years in respect of protein content. While the range of values obtained was greater in 2006, the overall mean was also significantly increased. These and other aspects of the wheat properties led to significant effects on baked product performance, particularly for Spiral white bread.

Significant variability in wheat and flour properties was also evident within populations. As an example, Figure 10 shows how milling extraction rate varied

within the Hereward×Malacca population for samples from both harvest years (2005 and 2006).

Figure 10. Milling extraction rate for Hereward×Malacca samples from 2005 and 2006 harvests



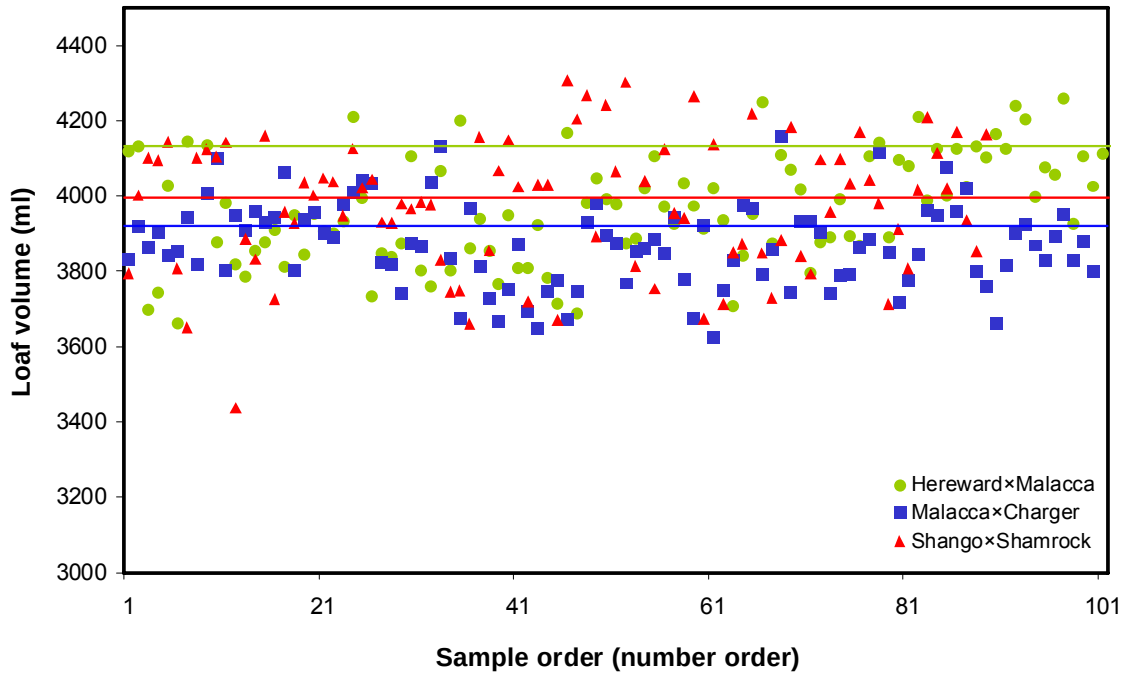
It can be seen that the number of samples processed in 2005 (88) was slightly lower than in 2006 (101). There were several lines in the first year which were not available for assessment but which were milled and baked in 2006. It should be noted that lines have been plotted by line number so results may be compared directly. It is clear that there was also a difference in the mean results between the two years with extraction rates being 0.8% higher on average in 2006. While this will be affected to a small extent by the difference in numbers, it appears that there is an underlying real effect of harvest year on extraction rate. This would tend to agree with the other results on wheat and flour quality where the 2006 harvest was generally seen as giving better performance. Although Figure 10 is not presented in a way which allows the correlation of results between years to be compared directly, it is clear that there are a number of areas where common patterns may be observed.

The first two points in both years represent the parents and it is clear that the population demonstrates transgressive segregation in both years with values being obtained for many lines which are either higher or lower than the results obtained for the parents. This clearly demonstrates the success of the population development stage with the lines selected for final assessment giving a good opportunity for the development of robust QTL. This pattern of results also indicates that there is scope for future improvements in wheat performance through breeding based on the fact that both parents make positive but different contributions to quality, in this case, milling performance.

3.4 Baked product quality

Given the range of wheat and flour quality parameters, it was not surprising to find that bakery performance varied significantly also. The full details may be found in the attached electronic database but Figure 11 shows an example of this variation for Spiral white loaf volume for samples from the 2006 harvest.

Figure 11. Loaf volume for Spiral white loaves for all populations from the 2006 harvest

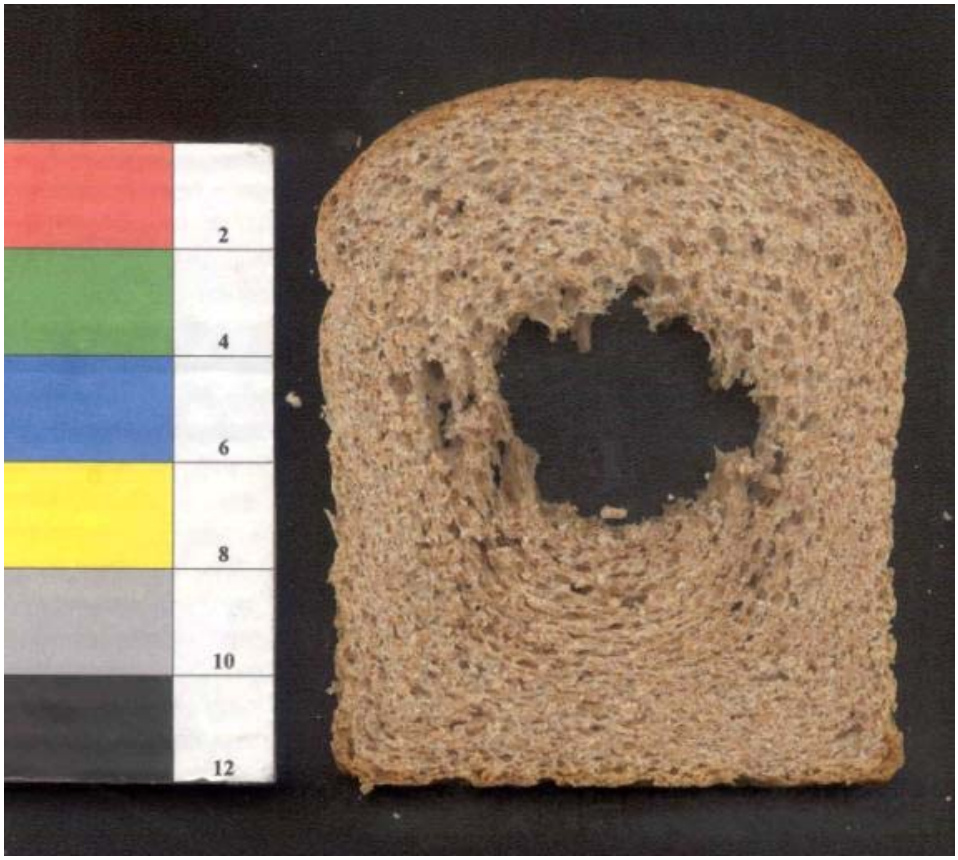


As with the wheat and flour results, there is evidence of significant variation between lines as well as some evidence of overall differences between populations within a

single harvest year. Transgressive segregation is also apparent and in this case, the parent line having the larger value has been illustrated using a line having the same colour as the relevant data series. This allows the lines within a population which yield superior volume potential for a given product (in this case Spiral white bread) to be clearly identified. For each population, it is important to note that a number of such samples were found. This confirms the selection process during population development as well as illustrating the potential for generating new varieties likely to have superior baking performance (as assessed by loaf volume).

In addition to the objective characters which were assessed of all bread products and puff pastry, overall observations of baking performance were also recorded where appropriate. One additional aspect of processing tolerance which resulted from this was the occurrence of holes within wholemeal products, apparently as a result of a collapse in the internal structure at the end of proof or during the early stages of baking (Figure 12).

Figure 12. Representative sample of wholemeal bread demonstrating the structural impact of hole formation



The sample in question emanated from the Malacca×Charger population and indeed all but one of the twelve samples showing such problems were from this population. The other sample was a line from the Shango×Shamrock population.

Interestingly, this phenomenon only occurred for samples from the 2005 harvest, indicating that there was an overall change in flour properties that exacerbated a propensity for such problems to occur for particular lines. Further assessment of processing dates and associated conditions did not lead to obvious reasons for the phenomenon's occurrence and so it appears that it is related to some aspect of the quality of the lines concerned and, as such, may be worthy of further evaluation in any further work.

To summarise the range of properties generated within the baking assessment and to enable the basis for sample selection to be visualised, representative images of selected lines for each product in each year have been included (Figures 13-20). It may be seen that the selection process has been effective in targeting lines in each population which cover a range of properties. Equally, it is clear that the processing approach also has a significant role to play in understanding the range of baked product characteristics which may be encountered. For the bread products, the volume potential and internal structure strongly depends on the processing route, with that generated using CBP showing particularly fine and uniform crumb structures. Equally while volume differences will still occur with lidded tins (as used in this case), differences in this parameter between the best and poorest loaves is less evident when comparing the results with those from the unlidded Spiral white bread.

The differences in performance between years is also marked, with the samples in 2005 generally showing significant variation in volume for Spiral white. Such differences were less marked in 2006 where average loaf volumes for this product were much higher. It would be expected, therefore, that the discrimination in this year between good and poor performing samples will be more affected by crumb structure, colour and firmness characteristics (all of which were also used in the sample ranking system).

Figure 13. Representative examples of white, CBP loaves from 2005 harvest

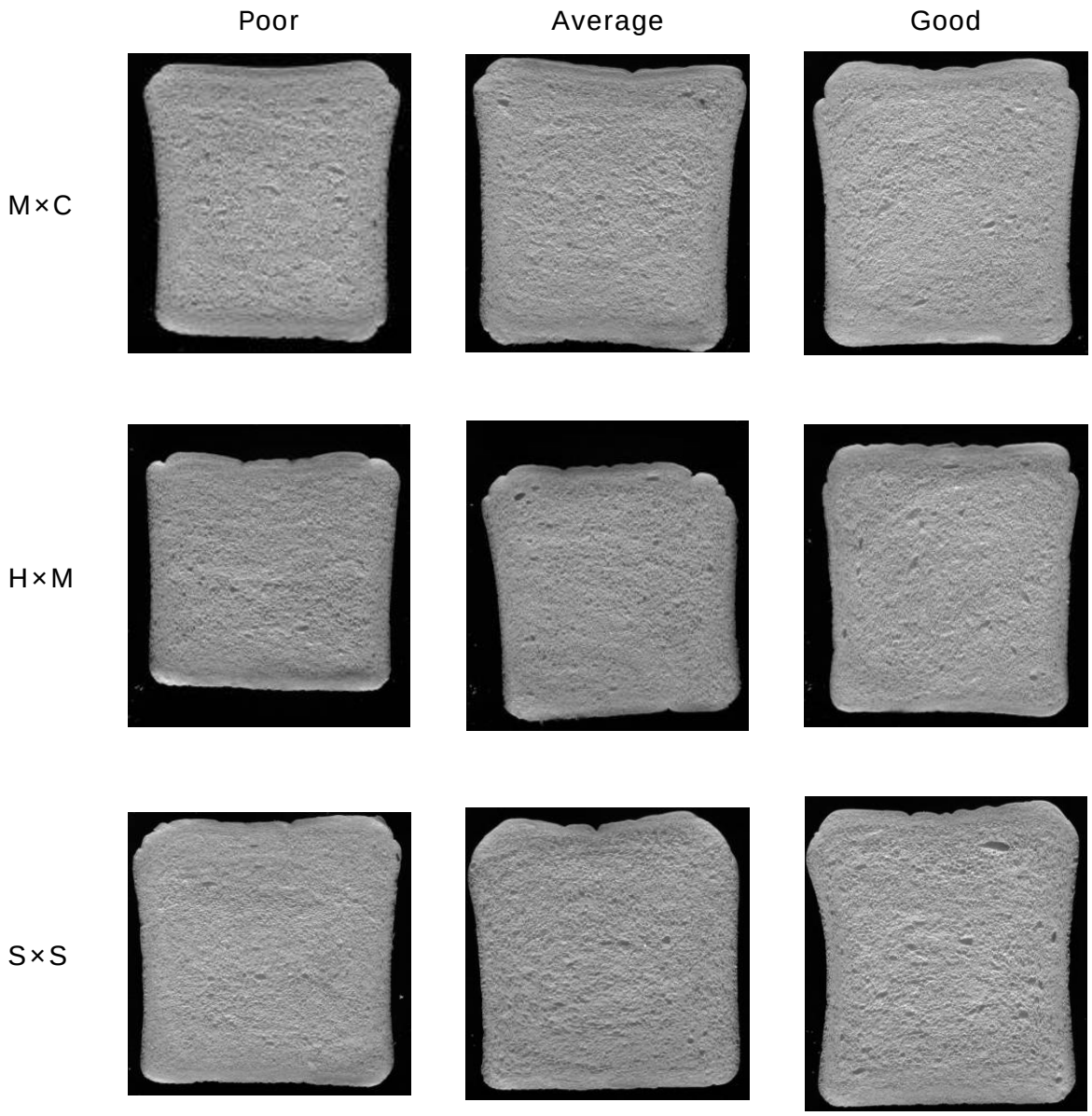


Figure 14. Representative examples of white, CBP loaves from 2006 harvest

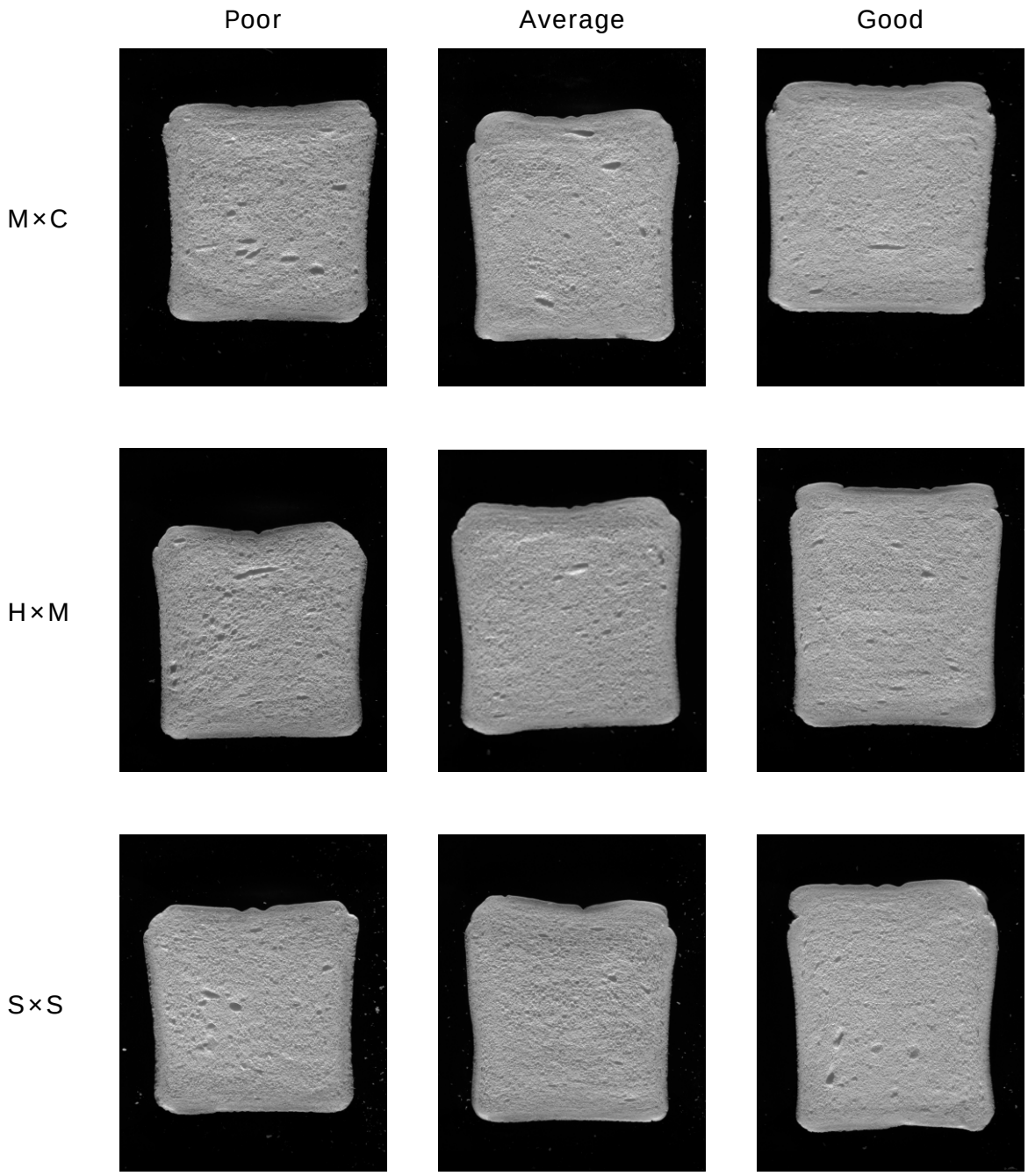


Figure 15. Representative examples of wholemeal loaves from 2005 harvest

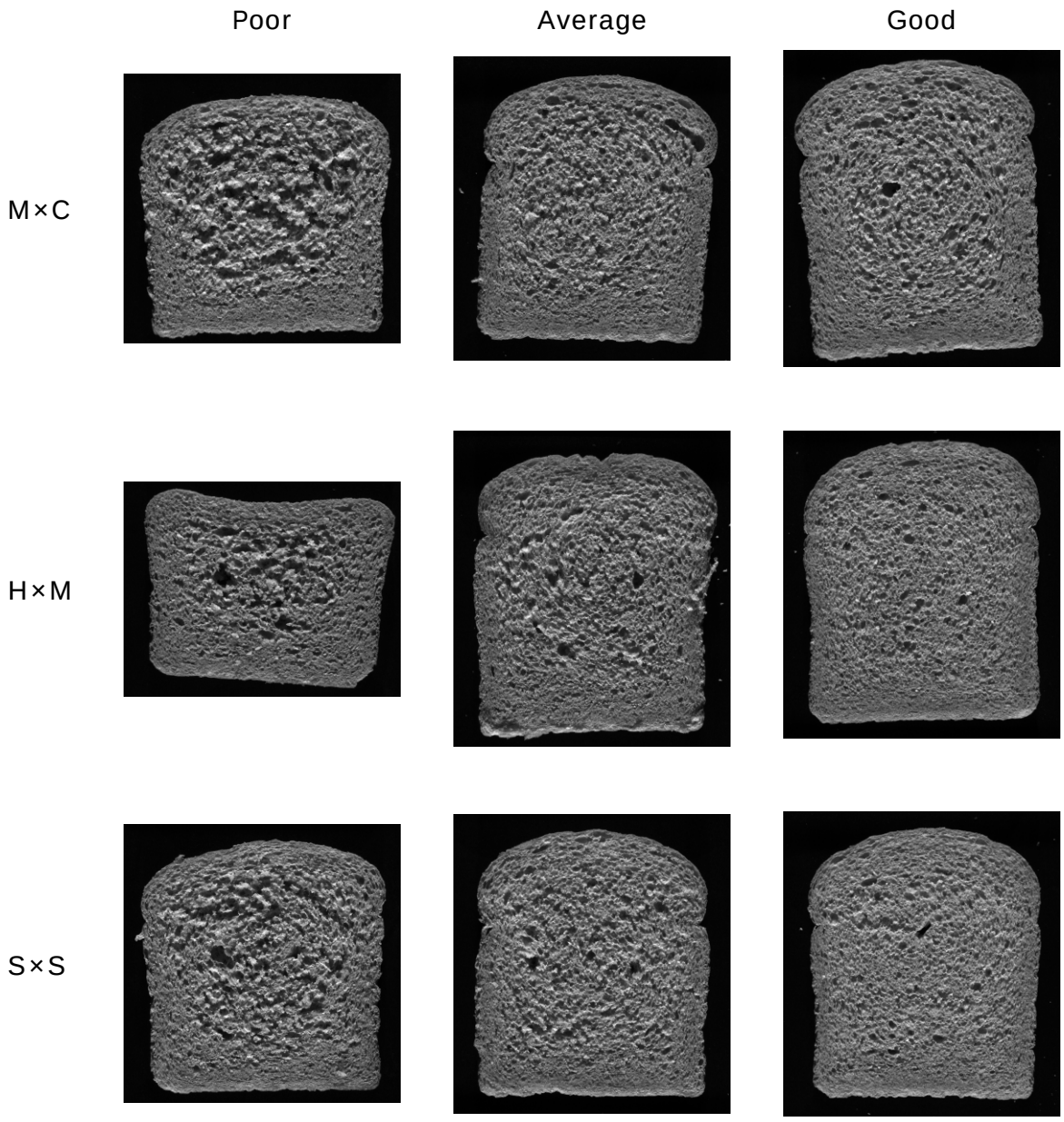


Figure 16. Representative examples of wholemeal loaves from 2006 harvest

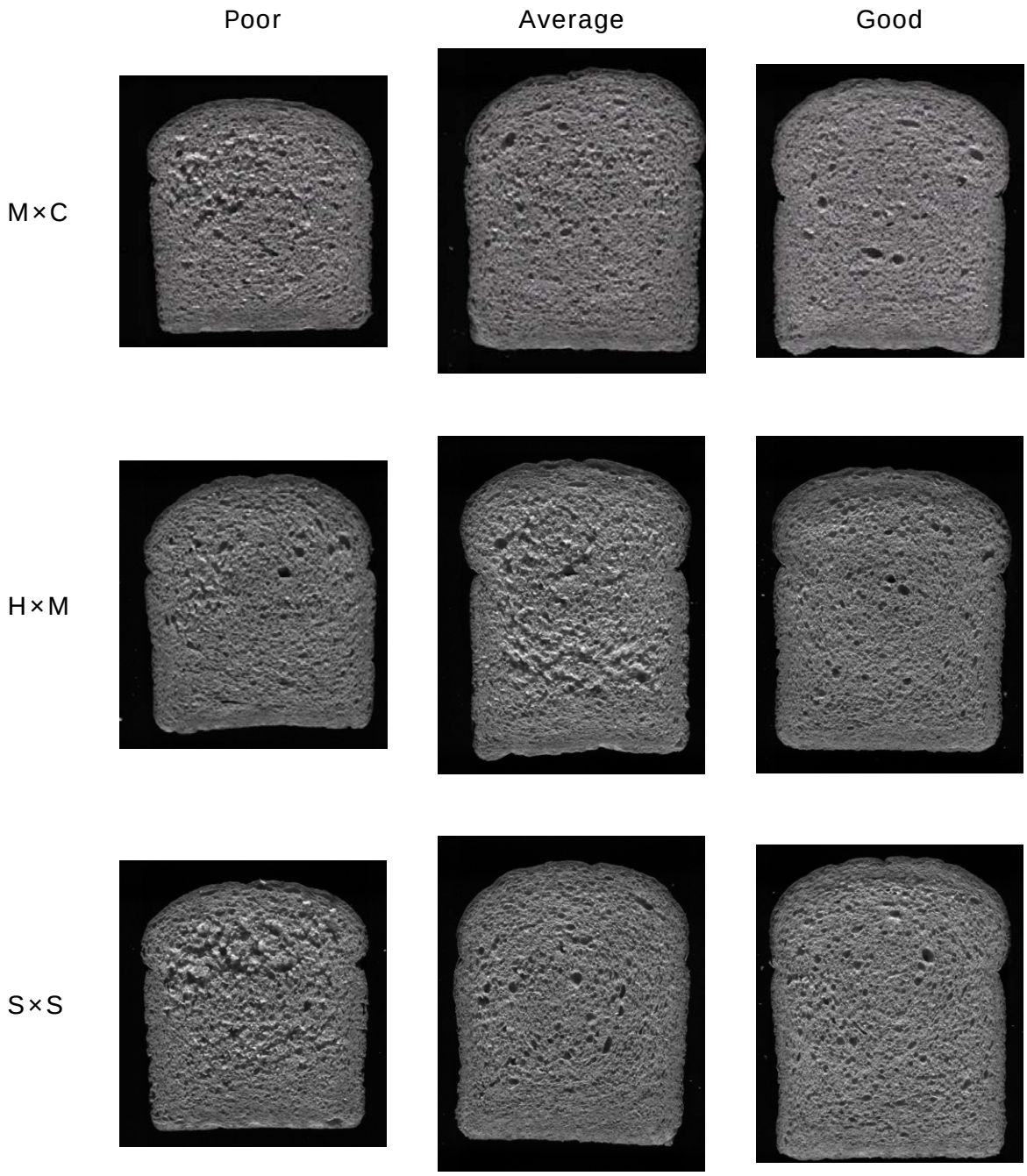


Figure 17. Representative examples of Spiral white loaves from 2005 harvest

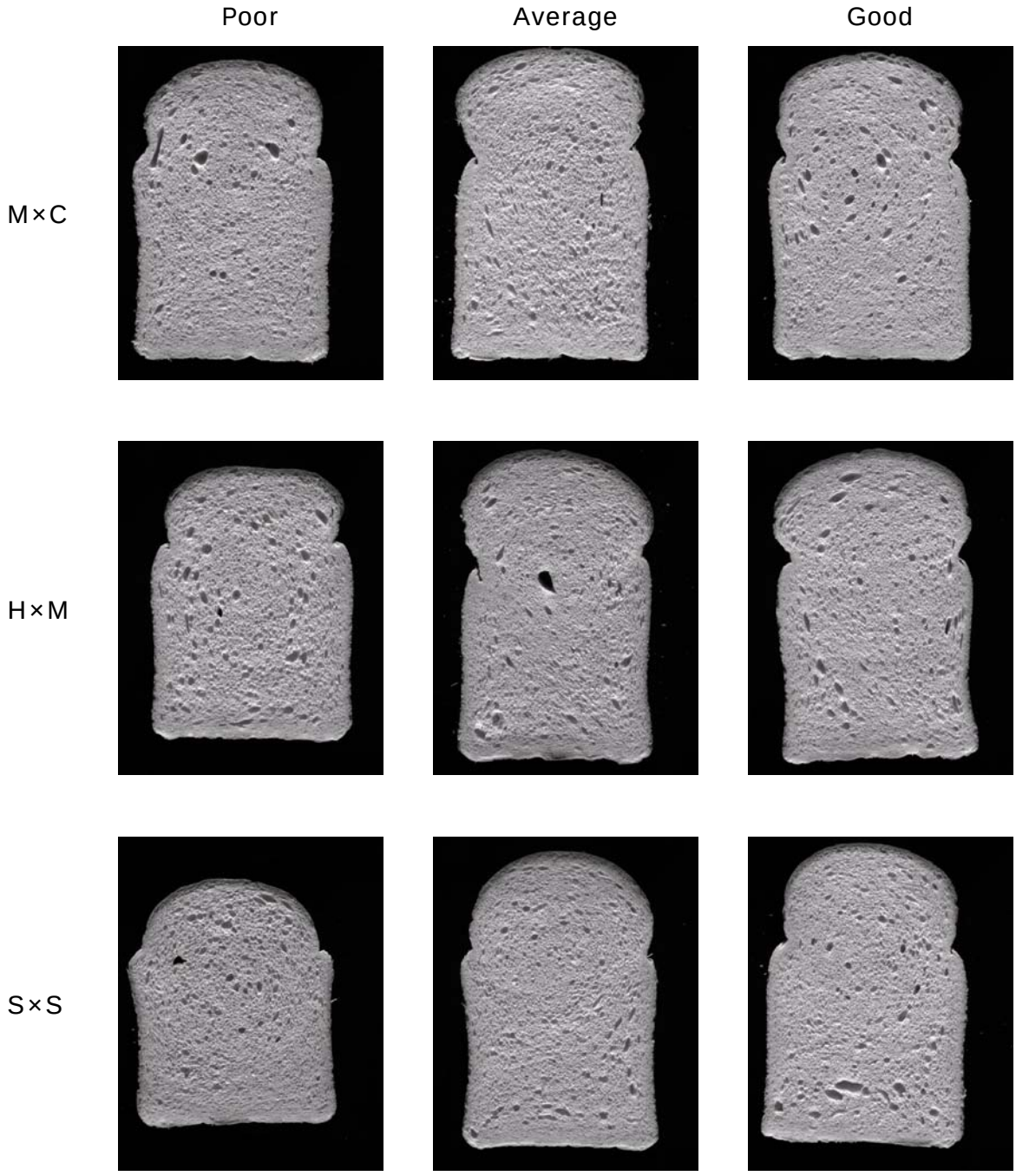


Figure 18. Representative examples of Spiral white loaves from 2006 harvest

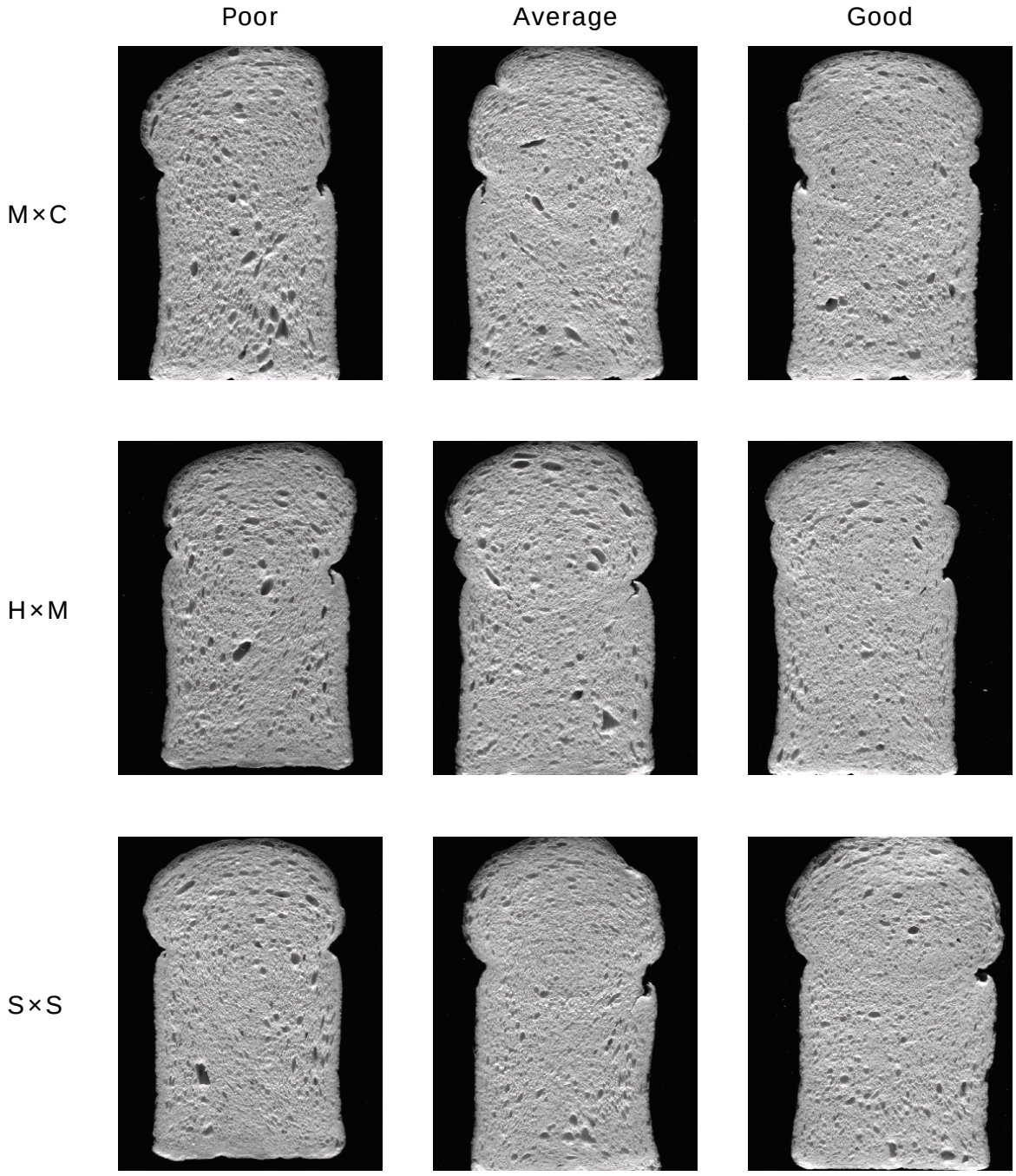


Figure 19. Representative examples of puff pastry from 2005 harvest

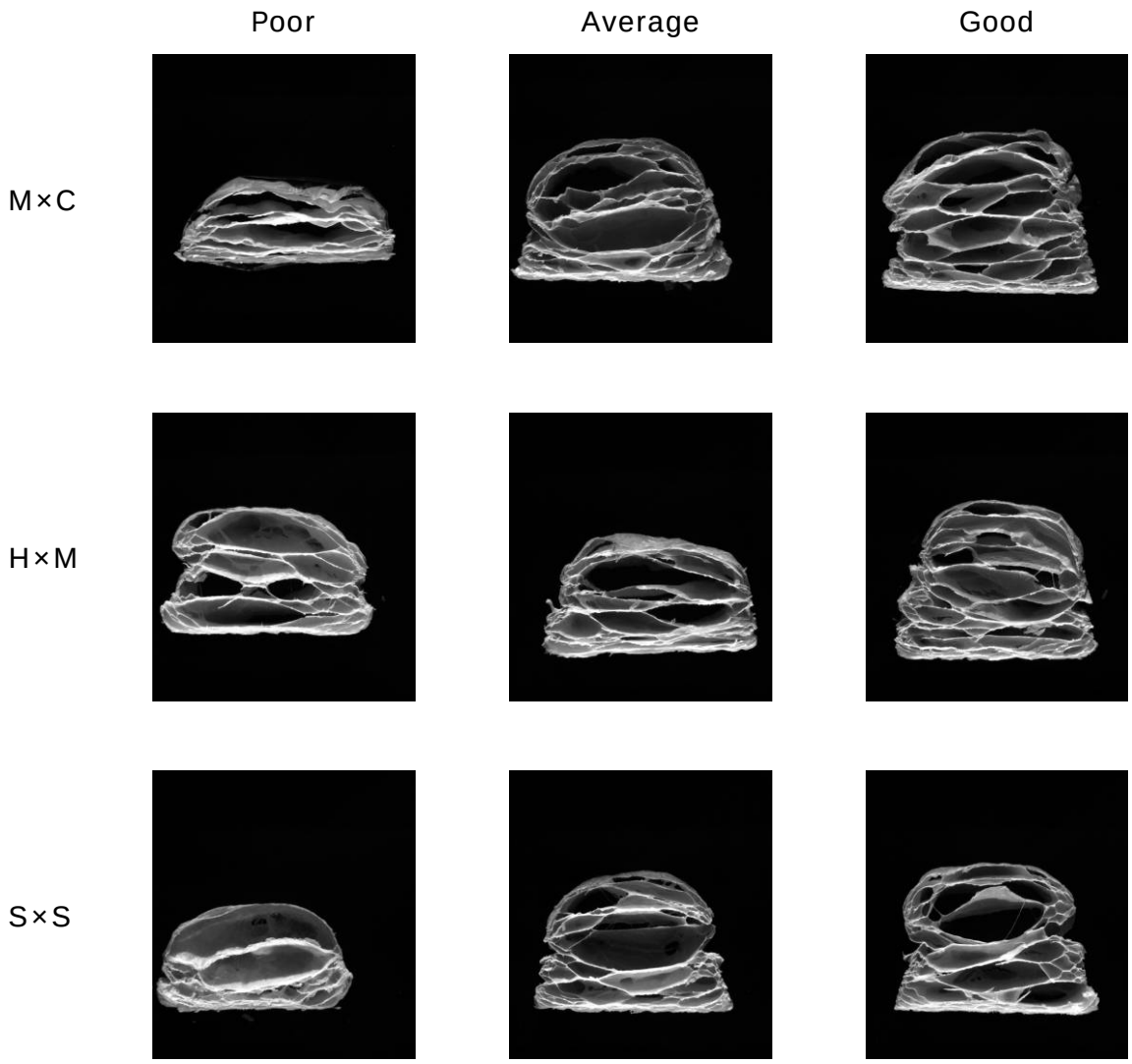
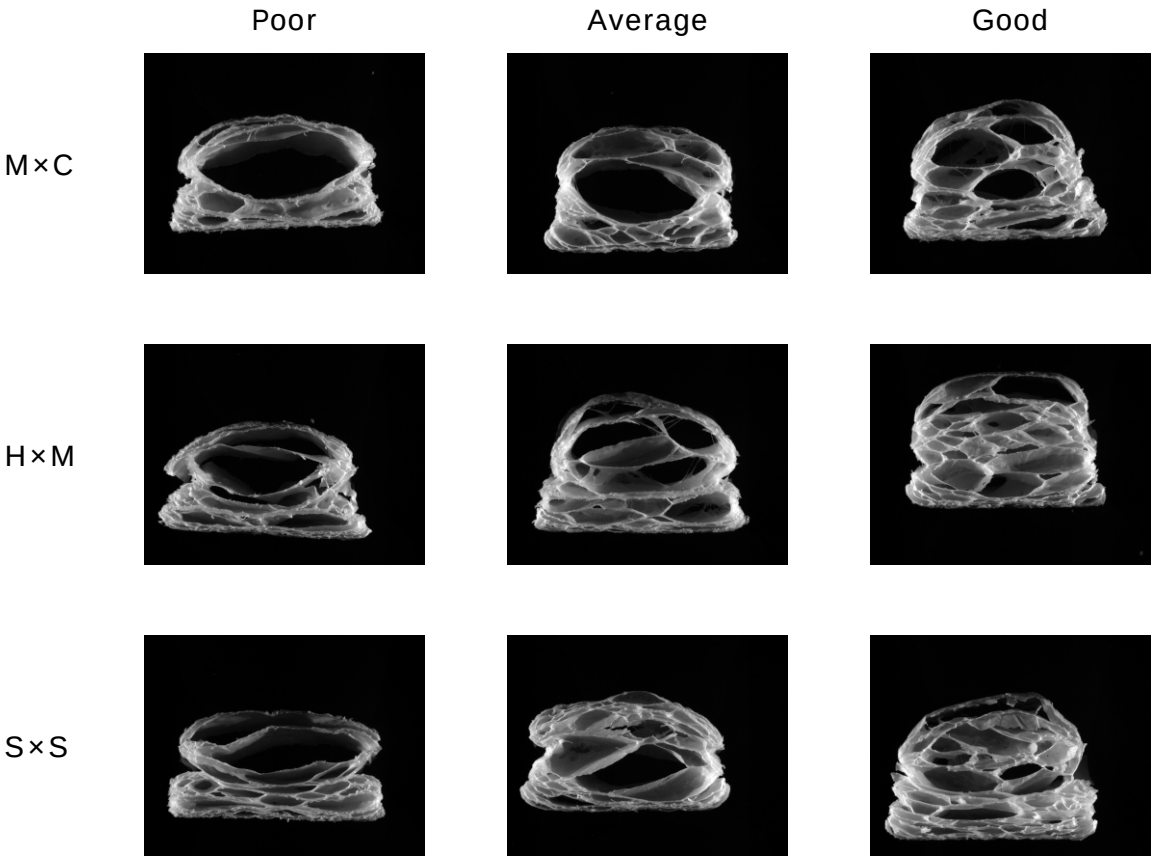


Figure 20. Representative examples of puff pastry from 2006 harvest



3.5 The development of the genetic maps of the three wheat crosses

Genetic maps of each of the populations were developed and are shown in Appendix 2. The detail of each of the three crosses is described below:

3.5.1 Malacca × Charger

For this population, 98 lines were mapped to correspond with those selected for end use analysis. A total of 579 SSR markers were screened for this cross, of which 246 were polymorphic. In addition to this, 213 reliable polymorphic DArT markers were provided, giving a total of 459 polymorphic markers to construct the map. Of these markers 120 were co-segregating and 25 could not be linked, leaving 274 loci mapping in to 44 linkage groups, covering a combined distance of 1696cM.

3.5.2 Hereward × Malacca

This mapping population contained 115 DH lines of which 15 were removed for containing non-parental alleles. 100 lines were used to create the genetic map.
In total, 580 SSR markers were screened, of which 249 were polymorphic. In addition to these, a further 305 markers were provided from DArT analysis, giving a total of 554 polymorphic markers to map. However, of these, 203 co-segregated and 26 had no linkage. The map consists of 288 markers mapping into 35 linkage groups, with a total distance of 1021cM.

3.5.3 Shamrock × Shango

Of the 76 initial DH lines for this population, 7 were removed as they contained non-parental background. This left 69 lines to be used for end use analysis. A further 18 lines developed at the JIC were also used to develop the map. A total of 489 SSR markers were screened for this population, with 215 being polymorphic. DArT analysis provided an extra 233 additional markers, for a total of 448 for map construction. There were 144 co-segregating and 26 with no linkage, giving a map of 263 loci in 35 linkage groups, with a total distance of 1337cM.

There are several monomorphic regions for each of the population maps where it was not possible to identify polymorphic markers. This suggests that the parental lines contain some genomic regions that are identical by descent.

3.6 The location of new QTL for end use quality

3.6.1 Correlation analysis

Most of the correlations were highly significant for major traits for the relationship between the quality performance across 2005 and 2006 within each individual cross and within each product (see Table 10 for the Hereward × Malacca cross). This is very important, as it indicates that the standardisation of the processing by the companies across the years and products has led to good quality data for QTL analysis. The poorest correlations were for CBP white traits, indicating that this product is more sensitive to process than to ingredient compared to wholemeal or Spiral white. However, it was gratifying that most correlations for loaf volume in the three crosses and three bread processes, apart from S×S Spiral white and M×C Spiral white and CBP white, were highly significant. Disappointingly, correlations between the puff pastry parameters over years were very poor and generally non-significant in all three crosses (Table 16).

Table 16. Correlation coefficients between years for puff pastry traits.

	M×C 2005-06		H×M 2005-06		S×S 2005-06	
	Cor.	Sig.	Cor.	Sig.	Cor.	Sig.
Avg Height (avg)/px	-0.015	NS	0.015	NS	0.007	NS
Avg Breadth / px	0.532	***	0.124	NS	0.019	NS
Avg Coarse /Fine Clustering	0.177	NS	0.007	NS	-0.155	NS
Avg No. of Cells	0.231	*	-0.013	NS	-0.007	NS
Avg Cell Volume	-0.100	NS	0.155	NS	0.175	NS
Avg Average Cell Elongation	0.051	NS	0.069	NS	-0.072	NS
Avg Cell Alignment	-0.104	NS	-0.038	NS	0.056	NS
Avg Non-Uniformity	-0.123	NS	-0.131	NS	0.161	NS
Avg Wall Thickness / px	-0.116	NS	-0.089	NS	-0.035	NS
Avg Count >10	0.005	NS	0.184	NS	0.064	NS
Avg Max Bubble px	-0.063	NS	0.002	NS	0.022	NS
Avg Calc Score	-0.177	NS	0.172	NS	-0.12	NS
Avg Tenderness	-0.087	NS	0.037	NS	0.239	*
Avg Layer Count	0.129	NS	0.095	NS	0.156	NS
Avg Height (mm)	0.089	NS	-0.004	NS	-0.131	NS
Avg Width (mm)	0.104	NS	0.153	NS	-0.024	NS
Avg Depth (mm)	0.280	**	0.052	NS	0.014	NS
Avg Calc % Shrinkage	0.214	*	0.212	NS	-0.016	NS

This was also reflected in the QTL analysis (see later). As the control flour performed in the same way in both years it is unlikely that this variability is due to processing. A possible cause is that for a given cross the protein difference from year to year was large enough to render it more or less suitable for the process employed. For instance, one of the best samples from the 2005 crop performed extremely poorly in 2006 due to the strengthening of its protein beyond the optimum for the method.

3.6.2 QTL analysis

Because of the huge datasets available, and hence the probability of finding false positives, the QTL analysis and interpretation focused on finding:

- QTL consistent over years.
- QTL consistent across characters and giving a 'story'.
- QTL found in the same location in different crosses.
- QTL found for the same characters in different products.

Where QTL are found to co-locate, this can either mean the pleiotropic effects of the same gene(s), or the effects of separate, but linked, genes. In many cases, knowledge of the breadmaking process can provide intelligent guesses of which is most likely to be occurring.

The very large number of traits analysed within each cross resulted in the discovery and localisation of many QTL for the different traits and products, most of which were new. This produced a UNIQUE DATASET of new and novel information on the genetic control of end use quality in wheat. No other groups, worldwide, have been able to produce this breadth and depth of information on the genetic control of breadmaking characteristics.

3.6.3 QTL analysis for milling performance

The QTL localisation data were summarised into Excel spreadsheets, and Table 17 gives an example of the summary data for the chromosomal location of QTL for milling traits in the Malacca x Charger cross over both years. As an example of new QTL found, it can be seen from Table 17 that there are highly significant QTL found on linkage group 14 controlling water absorption and development time, where the allele increasing water absorption comes from Charger. Other QTL controlling water absorption are found on linkage groups 22, 31 and 38.

Similarly, there is a large and significant QTL on linkage group 31 controlling Hagberg Falling Number, where the allele increasing Hagberg Falling Number comes from Malacca. This QTL clearly has 'knock-on' effects on dough stability. A range of other QTL was found in the other crosses, giving new insights into the allelic variation available to UK breeders for milling quality.

3.6.4 QTL analysis for puff pastry performance

QTL analysis of the puff pastry data from the 2005 harvest suggested that there were several new QTL involved in the genetic control of this product's performance. However, very few of these QTL were found again in the analysis of the 2006 harvest, and some of the QTL found even had opposite allelic effects over the seasons. Table 18 summarises the chromosomal locations of the QTL found in the Hereward × Malacca cross, indicating the inconsistency of results.

Table 17. Summary of Malacca × Charger milling QTL over both years

		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005
		9	8		**10**			**11**	9		**12**	10		**13**	11		**14**	12		**15**	13			
	Linkage Group	Water absorption (600 line), %	Water absorption (600 line), %		Water absorption (600 line), % (@14% moisture)			Development time, min	Development time, min		Stability , min.	Stability , min.		Degree of softening, BU	Degree of softening, BU		Falling No.(s)	Falling No.(s)		Water absorption (600 line), %	Water absorption (600 line), %			
	1																							
	2																							
	3	** M	** M		* M			* M	** M					* C	* C					** M	** M			
	4	** M			** M				* M		* C									** M				
	5								* C												** C			
	6																							
	7	* M			* M						*** C			*** M			* C							
	8	** C	* C		** C						** C	** C		** M	*** M			* C		** C	** C			
	9										* M			** C										
	10											** M		* C	** C		* C							
	11								* M					* C										
	12	* C			* C			* M																
	13							*** M	* M			** M									** M			
	14	*** C	*** C		*** C			*** C	*** C		* C	** C		* M	* M		** M	** M		*** C	*** C			
	15																							
	16	* C										* C					* C			* C				
	17													* M										
	18		* M												* C									
	19	* C			** C																			
	20																							
	21				* M							* C												
	22	*** M	*** M		*** M				** M											*** M	** M			
	23										* C			** M										
	24																							
	25																							
	26																	* M						
	27		** C								* M	** M		** C	** C						** C			
	28		* C		* C						* M			** C			* C	* C						
	29		** C																	** C				
	30							** M				** M			** C		* C							
	31	* M			* M			*** M	** M		* M	*** M			** C		*** M	*** M		*** M	** M			
	32	** C	*** C		*** C			* C	** C		* C	*** C			** M		* M	* M		** C	*** C			
	33																							
	34		** M					* M				** M			*** C						** M			
	35																							
	36	* C			* C																			
	37		* M						** M		* M			* C			** M				* M			
	38	** M	* M		** M			*** M	* M		* M	*** M		** C	** C					*** M	*** M			
	39										* M	* M		* C				* M						
	40										* C						* M	*** M						
	41		* M											* M						* M				
	42																							
	43										* M			* C				* M						
	44													* M				* C						
		9	8		10			11	9		12	10		13	11		14	12		15	13			

Table 18. Summary of puff pastry QTL found in the Hereward × Malacca cross over years

	2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005
	1	1		2	2		3	3		4	4		5	5		6	6		7	7		8	8			
linkage group	Avg Height (avg)/px	Avg Height (avg)/px		Avg Breadth / px	Avg Breadth / px		Avg Coarse /Fine Clustering	Avg Coarse /Fine Clustering		Avg No. of Cells	Avg No. of Cells		Avg Cell Volume	Avg Cell Volume		Avg Average Cell Elongation	Avg Average Cell Elongation		Avg Cell Alignment	Avg Cell Alignment		Avg Non-Uniformity	Avg Non-Uniformity			
1																	** M			** M		* M	* H			
2				* H							** H		*** M			*** M			* M	** M		** H				
3	* M	* H			** H			*** M						* M					** H							
4	* M									* M																
5										* M																
6														** M					* H							
7					* H						* H					* H	* M		* H				* M			
8	* H			** H						* H				* M								* M				
9													* H									** H				
10														* H					** H	** M						
11																							* M			
12							* H	* M					* H	* M			* H			* H						
13				* M						*** M										* H						
14																										
15		** M																				* H	* M			
16								** M		* H																
17																										
18				** M																						
19																							* H			
20				* M						* M																
21																										
22																										
23																			** H			* H				
24	* M						* H												* M				** H			
25								* H						* H						* M						
26										** H													* H			
27														* M												
28		* H																								
29		*** H		* H	* H			* M			** H									* M					
30											* M									* H						
31	* M				* H		* H			* M	* H									* M						
32					* H			* M		* H				* M												
33					*** H						* H					* M										
34	* H			** M							* M						** H		* M							
35																* H										
	1	1		2	2		3	3		4	4		5	5		6	6		7	7		8	8			

This analysis obviously reflects the lack of correlation between the data for all crosses over the seasons (Table 16). Some of these QTL may be false positives, particularly at low significance levels. However, others are likely to be harvest/process year dependent. Consequently, it appears that puff pastry performance is very much dependent on season and process and, unfortunately, no clear guide to which alleles can be incorporated into new germplasm to give consistent performance over seasons can be gleaned from this analysis.

3.6.5 QTL analysis for CBP, wholemeal and Spiral white bread performance

In contrast to the puff pastry performance, there were highly significant and consistent QTL found for all the bread making processes, CBP white, wholemeal and Spiral white. As an example, Table 19 shows the chromosomal locations of QTL for CBP white in the Shamrock × Shango cross. Some of the QTL can be interpreted based on previous knowledge. For example, the QTL on linkage group 2 affecting a range of traits is most probably due to known allelic variation at the storage protein

loci known to be located on this chromosome. However, others, for example those relating to C-Cell measured parameters, such as the number of bubble cells in a loaf slice, are completely new. With respect to this parameter, it can be seen that consistent and significant QTL are found on linkage groups 5, 6 and 32, where the increasing alleles come from Shango, whilst linkage group 27 has an increasing QTL from Shamrock. The dispersion of increasing alleles between the parents has obvious breeding implications, indicating that transgressive segregants are possible for this trait. A similar situation is found for several of the other important traits relating to bread performance in the other crosses.

Table 19. Summary of CBP white bread QTL found in the Shamrock × Shango cross

	2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005		2006	2005
linkage group	Loaf volume, ml	Loaf volume, ml		L*	*,		Firmness @25% compression (Day 1), N	Firmness @25% compression (Day 1), N		Firmness @25% compression (Day 3), N	Firmness @25% compression (Day 3), N		Average Height_Max	Average Height_Max		Number_ of_Cells	Number_ of_Cells		Wall_ Thickness	Wall_ Thickness		Cell_ Diameter	Cell_ Diameter
1				** G	* G			** G			* G		* R	** R			* RG		* R	* R		* R	** R
2	* G	** G		*** G	** G		* R			* R			*** G	*** G		*** G	*** G			*** R			*** R
3											* G									* R			* R
4							* R						* G	* G		** G	* G		** R	* R		** R	* R
5										* R	* R						* G			* R			* R
6					** R		* G			* G				* G		** G	* G		* R			* R	* R
7		* R		* G			*** G			** G	* G						* R		** G	*** G		* G	*** G
8																			*** R			** R	
9													* G				* G						
10					* G																		
11																							
12																							
13																* G							
14							* G			* G													
15														* G									
16																						* R	
17																							
18																			** G			*** G	
19		* G			* G			** R		* R	** R			** G			* G		* G				
20		** R			* R								*** R	* RG			* R					** R	
21																							
22				** G				* R			* R			* G		* G							
23		** G												* G									
24																	* R			* G			* G
25		** G			*** G									* G			** G						
26		* G		* R									* G			* G			* R			* R	
27					* R		** G			* G			** R	** R		** R	*** R			** G			* G
28		** R			* R									** R			* R						
29		** G						* R			* R		* G	** G									
30				* R			* G			* G							* R						* G
31		** R			* R		* G			* G				** R			** R			* G			* G
32				* G	* G											** G	* G		** R	* R		* R	
33				* G												* G							
34																							
35					* R																		

It is of interest to compare if the same QTL or different QTL are segregating in the different crosses. This can be evaluated by comparing the QTL found for a particular

trait and product across the three crosses. Table 20 shows the QTL found for one of the most important traits, loaf volume, measured for CBP white bread across years and crosses, where the linkage groups are aligned homologously. Some strong QTL are segregating in only one cross, for example on linkage group 34 in the Hereward × Malacca cross where Malacca has a strong allele increasing volume relative to Hereward. But several others are consistent across years and crosses, for example the QTL on homologous linkage groups 9 and 7 in the Malacca × Charger and Hereward × Malacca crosses, respectively. Since these crosses have the Malacca parent in common, and the increasing allele comes from the other parent, this implies that Malacca has a 'weak' allele for loaf volume on the chromosome on which these linkage groups are found, which if replaced could increase its breadmaking quality.

Table 20. A comparison of loaf volume QTL found for CBP white bread across the three crosses and two harvest years, aligned by homologous linkage groups

Malacca x Charger			Malacca x Hereward			Shamrock x Shango		
	2006	2005		2006	2005		2006	2005
linkage group	Loaf volume, ml	Loaf volume, ml	linkage group	Loaf volume, ml	Loaf volume, ml	linkage group	Loaf volume, ml	Loaf volume, ml
3	* C		2	* M	* H	2	* G	** G
9	* C	* C	7	** H	* H	6		
10	* M	* C	8		** H	7		* R
17		* C	13	* M	* M	12		
28	* M	* C	25	** M		23		** G
31	* C		28			25		* G
32	*** M	* M	29		* H	26		
37	* C	** C	32			29		
38		* C	33			30		** R
39		* C	34	** M	** M	31		
40			35	* H		32		

The breadmaking traits for the different products were measured by different processors, and it could be that different genes interact with the different processes, and may represent different biochemical reactions in the dough. Thus, it is of

particular interest to see if this is the case and whether the same QTL control the same named traits across the products in the different crosses. Table 21 shows the QTL discovered for the trait 'loaf volume' measured following the three bread processes in the Hereward × Malacca cross. It is clear that there is quite a lot of consistency across products, particularly between CBP white and wholemeal. For example, the QTL for loaf volume on linkage group 13 is consistent across all product types and years, with Malacca possessing an allele which increases loaf volume relative to the alleles in Hereward.

Table 21. A comparison of loaf volume QTL found in the different products for the Hereward × Malacca cross

	Wholemeal		CBP White		Spiral white	
	2006	2005	2006	2005	2006	2005
linkage group	Loaf volume, ml	Loaf volume, ml	Loaf volume, ml	Loaf volume, ml	Loaf volume (ml)	Loaf volume (ml)
1						
2		* H	* M	* H		
3	** H	* H	* M		*** H	
4						
5	* M					
6						
7	* H	*** H	** H	* H		
8				** H	* H	** M
9					* H	*** H
10	* H					
11					** H	
12	* HM	* H				
13	* M	*** M	* M	* M	*** M	
14						
15			* H			
16						
17			** H			
18						
19		* H	* H	** H	** M	
20						
21						
22						
23						
24			* M			
25			** M			
26						
27		** H		* H		
28						
29	*** H			* H		
30					* H	
31	** H	** H				
32	* H	** H				
33					*** H	* H
34		* M	** M	** M		
35			* H		* M	* M

3.6.6 Co-location of QTL for different traits, years and products

As mentioned above, it is particularly informative to look at the intra-chromosomal location of QTL for all traits, products and years within a cross to look for the co-location of QTL for different traits, which will imply either the pleiotropic actions of single genes, or the linkage of multiple genes. To visualise this, the outputs for the locations of QTL from QTL Café from Interval Mapping were used to align the QTL with the chromosome maps for all important traits over products and years. As an example, Figure 21 shows the co-location of QTL for chromosome 2B in the Malacca x Charger cross for the range of milling and baking traits measured for the different products.

Several observations can be made. First, this is a 'QTL rich' chromosome with allelic variation for a number of traits dispersed along the whole length of the chromosome. Secondly, QTL for the same trait over years locate at the same place (within sampling error), indicating the fidelity of the original process and data on individual lines over years. Thirdly, there may be more than one QTL for a particular trait segregating on a chromosome. Again, for example, loaf volume is controlled by at least two QTL, one at the top of the short arm and one at the bottom of the long arm. Fourthly, there is probably a functional relationship between QTL. For example, the QTL for increasing loaf volume both at the top and bottom of the chromosome co-locate with QTL for more bubble cells. Thus bigger cells result in a bigger loaf, even though there is also a QTL for lower cell number co-locating. Thus a hypothesis for the relationship would suggest one QTL allele from Malacca at this location which increases bubble cell number, resulting in a bigger loaf with more small cells. Several other chromosomes are 'QTL rich' over the different products and from these comparative analyses, many other hypotheses can be formulated from the different types of QTL co-location to explain the relationships between the trait QTL discovered and the way in which product quality is controlled.

3.6.7 Comparisons between crosses

Because the same set of molecular markers were used to map the different crosses, it is possible to compare the QTL discovered within each cross across the three crosses by aligning the individual chromosomes maps using common polymorphic markers. As an example of this analysis, Figure 22 shows the alignment of QTL for chromosome 3A between the Hereward × Malacca and Malacca x Charger crosses. Three polymorphic markers, *Wpt 7341* (DArT marker), *gwm 369* and *wmc 169* (SSR markers) align the whole length of the chromosome.

Figure 21. Co-location of QTL on chromosome 2B in the Malacca × Charger cross

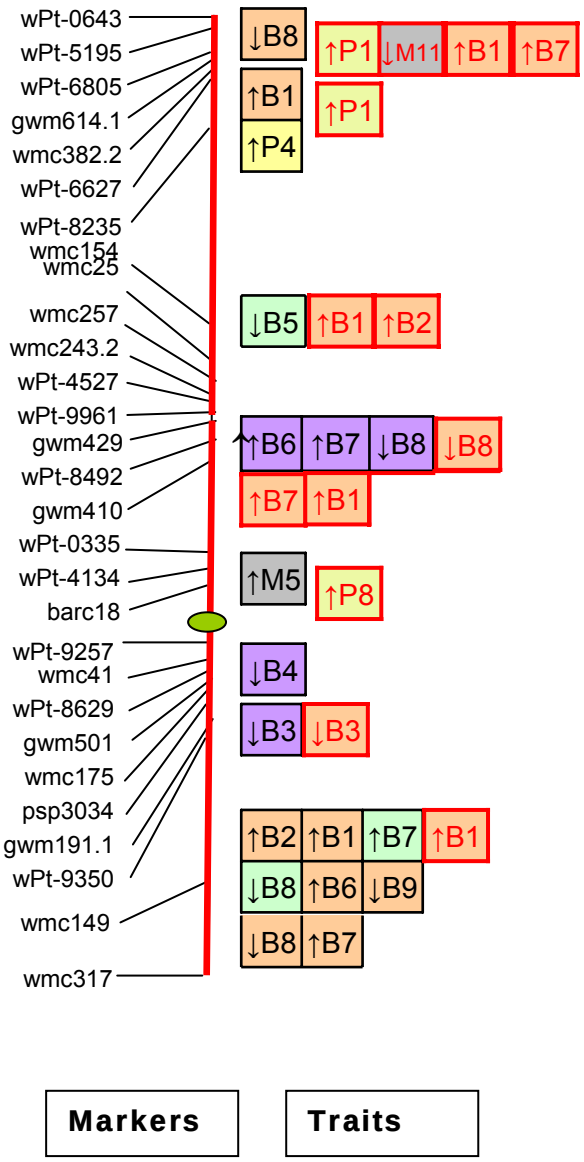
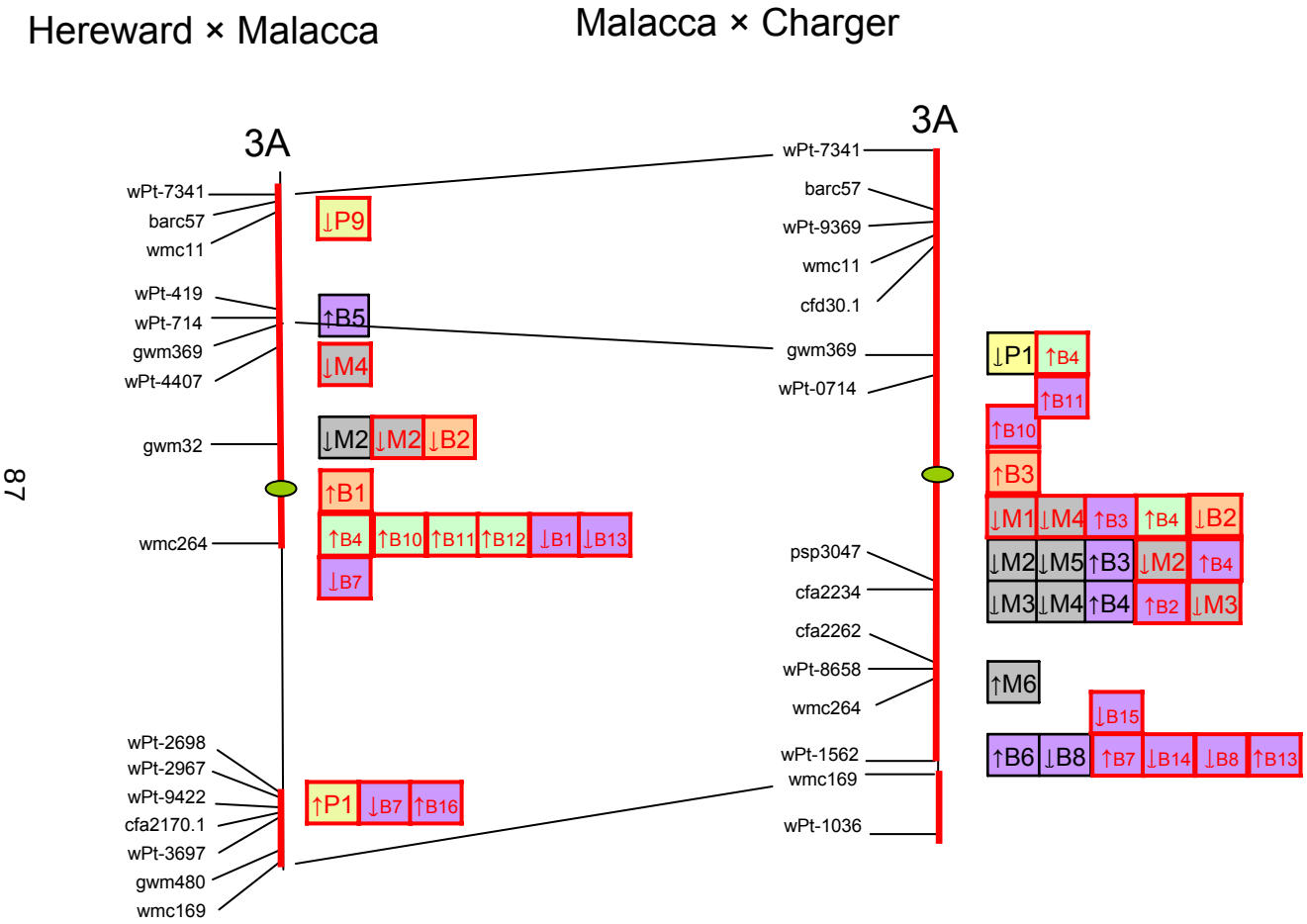


Figure 22. Comparative mapping for QTL on chromosome 3A between the Hereward × Malacca and Malacca × Charger crosses



In both crosses, there is a group of QTL for milling and baking traits that co-locate near the centromere in each cross and can be aligned across crosses. In both crosses and over years, a major QTL for protein content locates in this region, and this probably has a pleiotropic effect on the baking traits. These relationships can be found for all products, and give an excellent insight into the functional relationship between different traits and the allelic variation segregating in UK germplasm.

3.7 Metabolomics

3.7.1 PCA analysis

PCA is a data visualisation method that is useful for observing groupings within multivariate data. Data are represented in *n* dimensional space, where *n* is the number of variables, and is reduced into a few principal components, which are descriptive dimensions that describe the maximum variation within the data. The principal components can be displayed in a graphical fashion as a "scores" plot. This plot is useful for observing any groupings in the data set and in addition will highlight outliers that may be due to errors in sample preparation or instrumentation parameters etc. PCA models were constructed using all the samples in the study. Coefficients by which the original variables must be multiplied to obtain the PC are called "loadings." The numerical value of a loading of a given variable on a PC shows how much the variable has in common with that component. Thus, "loading plots" can be used to detect the spectral areas responsible for the separation in the data. The numerical value of a loading of a given variable on a PC indicates how much the variable has in common with that component.

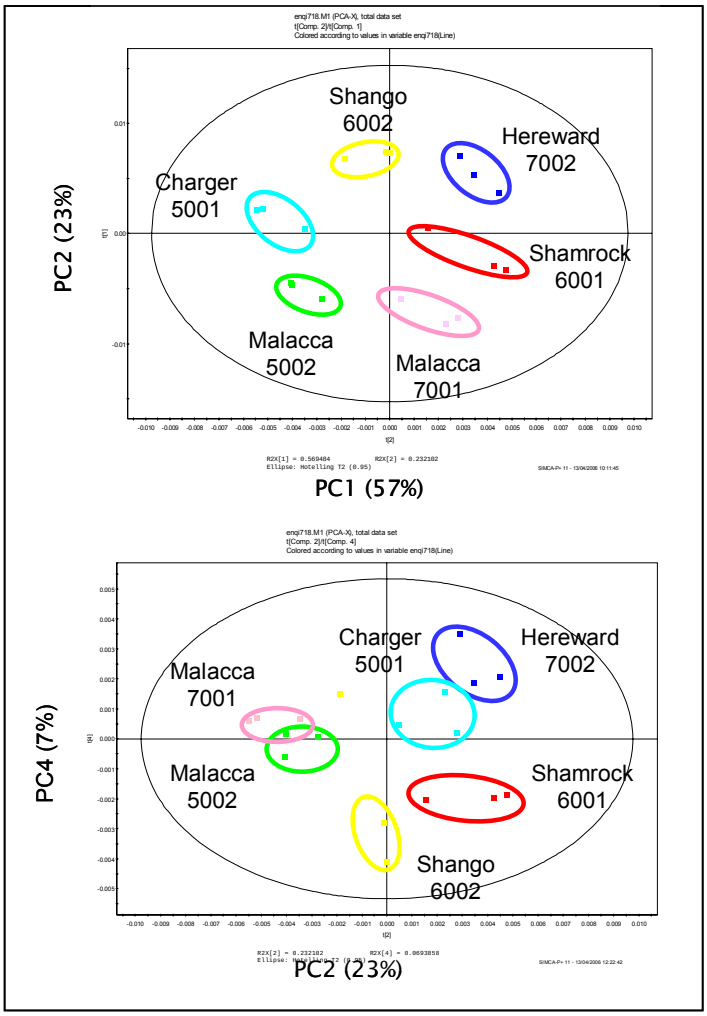
3.7.2 Analysis of parents

A preliminary experiment was carried out on the parent lines only. NMR data were collected and analysed, using PCA in SIMCA-P. As can be seen from Figure 23 each of the parent lines formed a distinct cluster when viewing the first 2 components of the statistical model (accounting for 80% of the total variance).

Interestingly, the clusters for the two Malacca samples could be separated from each other. These samples have been submitted as part of different populations and possibly reflect differences in sample handling at two different sites or on different occasions. Despite this separation, the Malacca samples resided in the same area of the scores plot relative to the other samples in the experiment. By the examination of

later components in the statistical model (e.g. PC2 vs PC4) the Malacca samples could be made to cluster together.

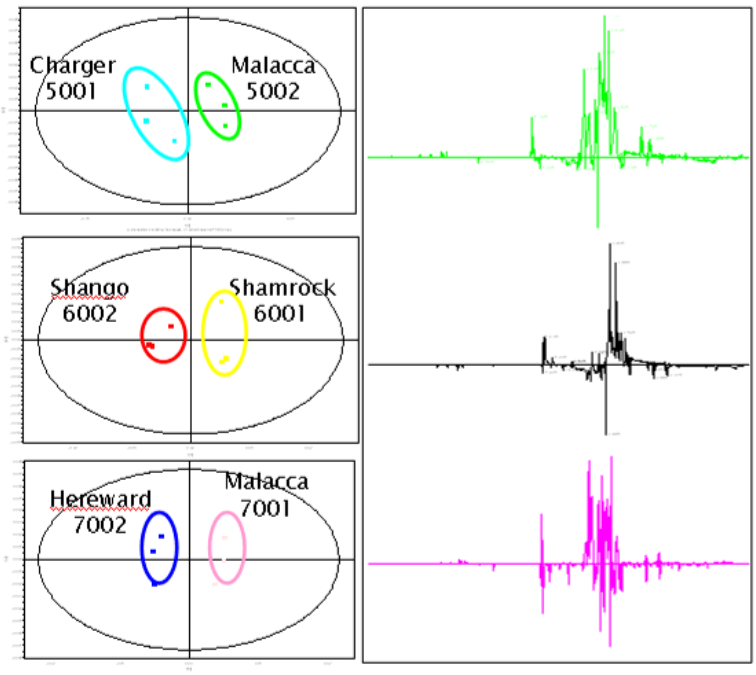
Figure 23. PCA analysis of NMR data from parent lines



The loadings plots generated from the PCA model help to describe the metabolite changes responsible for the separation in the data. When displayed as a line plot, these resemble the original NMR spectrum and thus can be compared to a library of standard compounds run under the same conditions. In order to fully explain the differences between the parents of each population, pairwise PCA models and resultant loadings plots were generated and compared (Figure 24).

The dominant features of the PCA loadings plots were carbohydrate signals. Malacca was found to contain increased sucrose, raffinose and aspartate when compared to Charger. In addition small increases were observed in the aromatic region.

Figure 24. Pairwise PCA models of NMR data from parent lines. Scores and resulting loadings plots



Compared with Shango, the Shamrock samples contained increased sucrose and maltose but decreased levels of raffinose and aspartate. Small decreases were also observed in the aromatic region. Hereward contained increased sucrose and maltose but decreased levels of an unidentified carbohydrate compound.

3.7.3 Analysis of populations

Three separate populations were submitted for analysis. Samples were submitted in two batches over two years. The initial batch consisted of 12 lines Malacca × Charger, 10 lines Shango × Shamrock and 86 lines Malacca × Hereward. In addition the parents for each population were also supplied for analysis. The following batch consisted of 12 lines Malacca × Charger, 11 lines Shamrock × Shango and 9 lines Hereward × Malacca. In addition the parents for each population were again supplied for analysis. Three analytical replicates of each biological sample were made for quality assurance purposes.

As the samples were supplied (and grown) in two separate batches the data from each batch were modelled separately (Figure 25) and then modelled together (Figure 26). The data were coloured according to population.

Figure 25. Analysis of 2005 (Batch 1) and 2006 (Batch 2) data

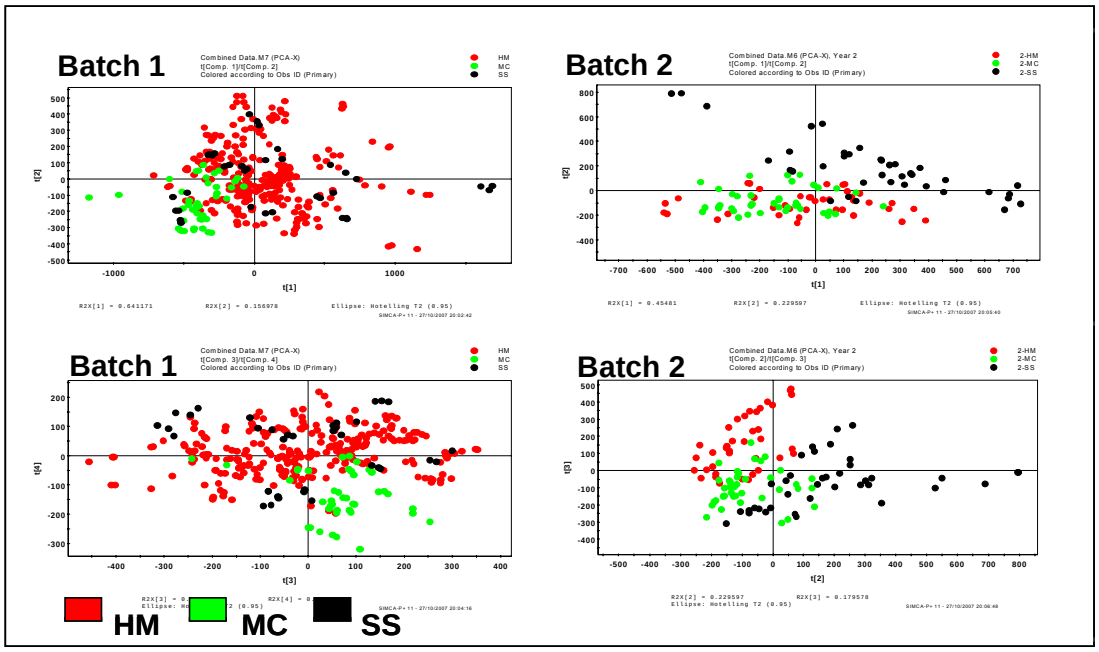
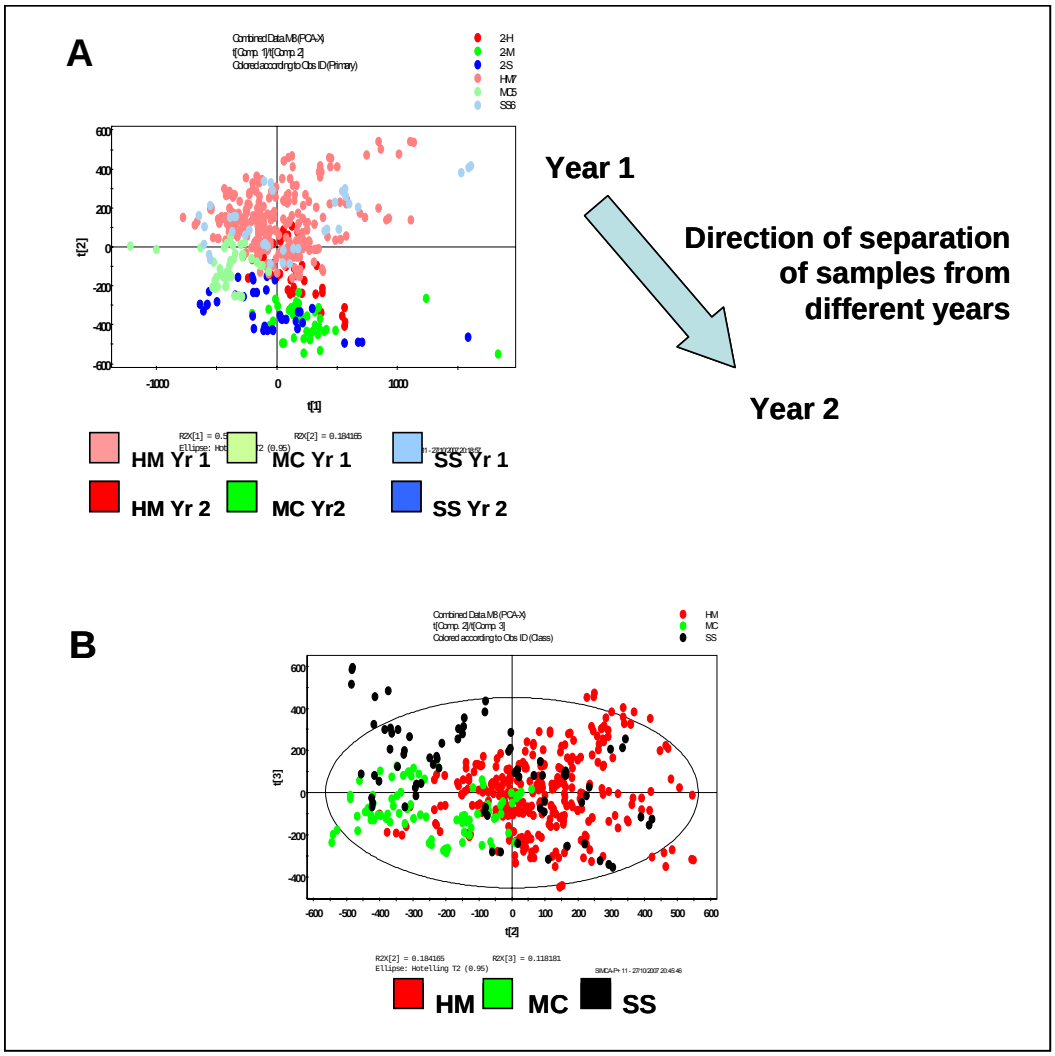


Figure 26. PCA Analysis of combined 2005 (Batch 1) and 2006 (Batch 2) samples. A- PC1 vs PC2 B- PC2 vs PC3, coloured by population



When the data from both years are combined it is clear to see that there is a separation by year when the first two principal components are analysed. Analysis of PC2 vs PC3 and colouring by population removes the year-to-year separation and the data cluster by population irrespective of the year of generation.

3.7.4 Correlation of NMR data with breadmaking quality data

The aim of the experiment has been to determine whether any metabolites in a global NMR profile of flour samples could be correlated against breadmaking quality data. In order to test this loaf volume, protein content and firmness data were appended to the NMR dataset and used in turn as categorical Y data in an orthogonal partial least squares (OPLS) model. OPLS is available to use when there is 1 response variable (Y) to analyse against the dataset. OPLS analysis separates the PLS components into two groups: components that are related to Y, called predictive and components that are orthogonal to Y called orthogonal. With one Y, there is one predictive component and several orthogonal (no relation to Y).

With every orthogonal component, a rotation occurs resulting in a certain percent of Y being explained. After all the significant components are extracted, the total percent of Y explained due to the rotation of the orthogonal components is computed.

The resulting Figure 27 is a scores plot which can be coloured from “low” to “high” for the selected Y variable. Correlations have been explored for the first and second batch of Hereward x Malacca samples and results were in agreement in both cases. The examples shown here are for the second smaller Hereward × Malacca batch of samples.

On initial inspection of the OPLS model, the metabolites responsible for the correlation appeared to be carbohydrate in nature. To further examine these chemical shifts and to validate their significance and correlative properties, the data were imported into Spotfire Decisionsite software which allowed column relationships to be explored via ANOVA. Example plots showing chemical shifts with both positive and negative correlations to loaf volume were easily identified and examples of such chemical shifts are shown below (Figure 28).

Figure 27. OPLS analysis of 2006 (Batch 2) H×M samples. Coloured by loaf volume

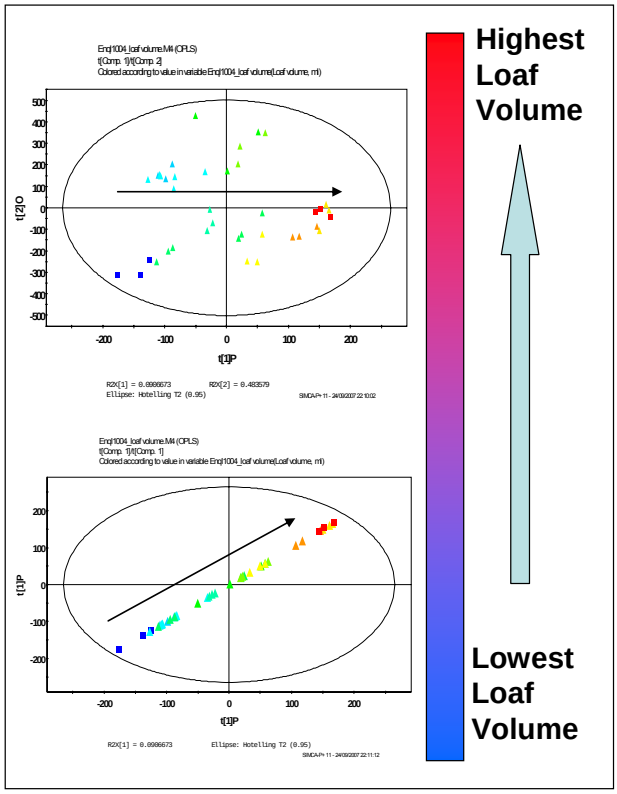
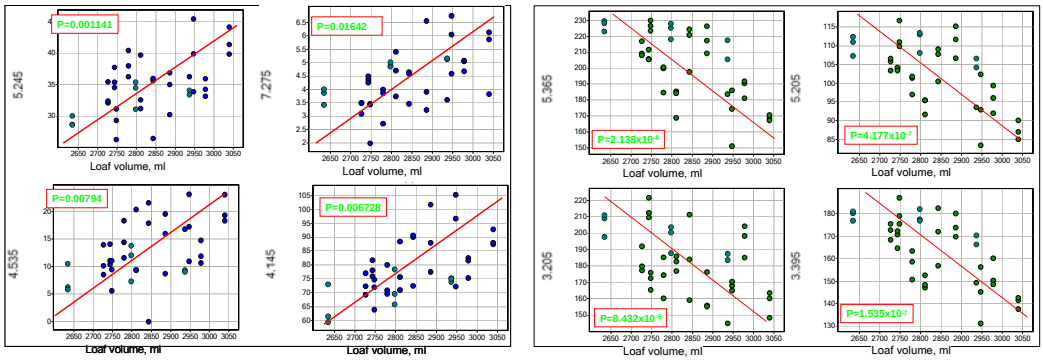
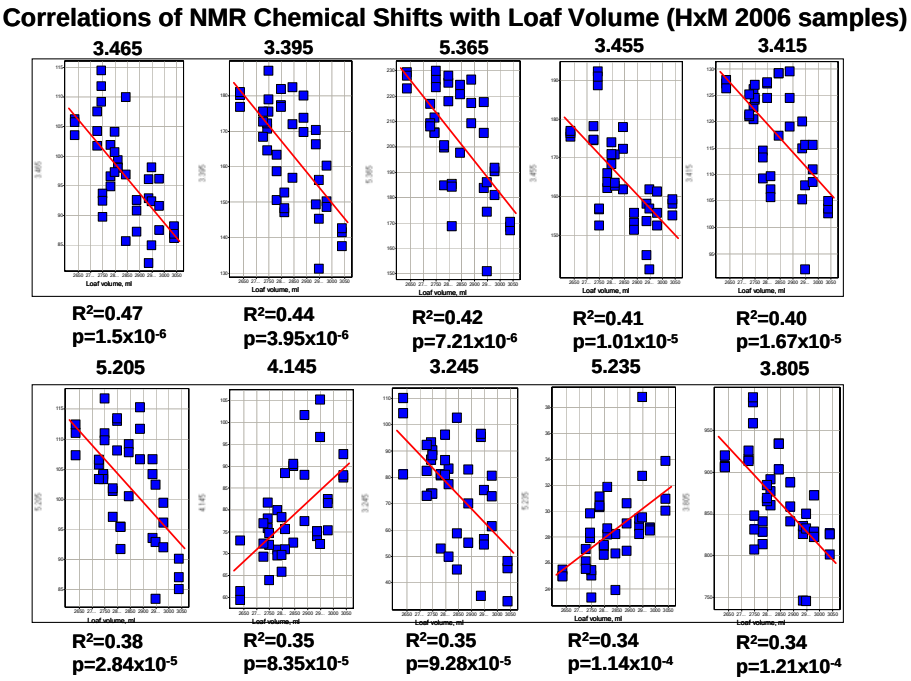


Figure 28. Examples of chemical shifts showing positive and negative correlations to loaf volume



Looking across the list of chemical shifts showing a significant p-value (Figure 29) it can be seen that of the top ten chemical shifts correlating with loaf volume, eight show a negative correlation with loaf volume, whereas only 2 show a positive correlation. R^2 values for these correlations are lower than one would usually accept (0.35 to 0.48); therefore it should be concluded that while these NMR chemical shifts do appear to correlate, this correlation is fairly weak.

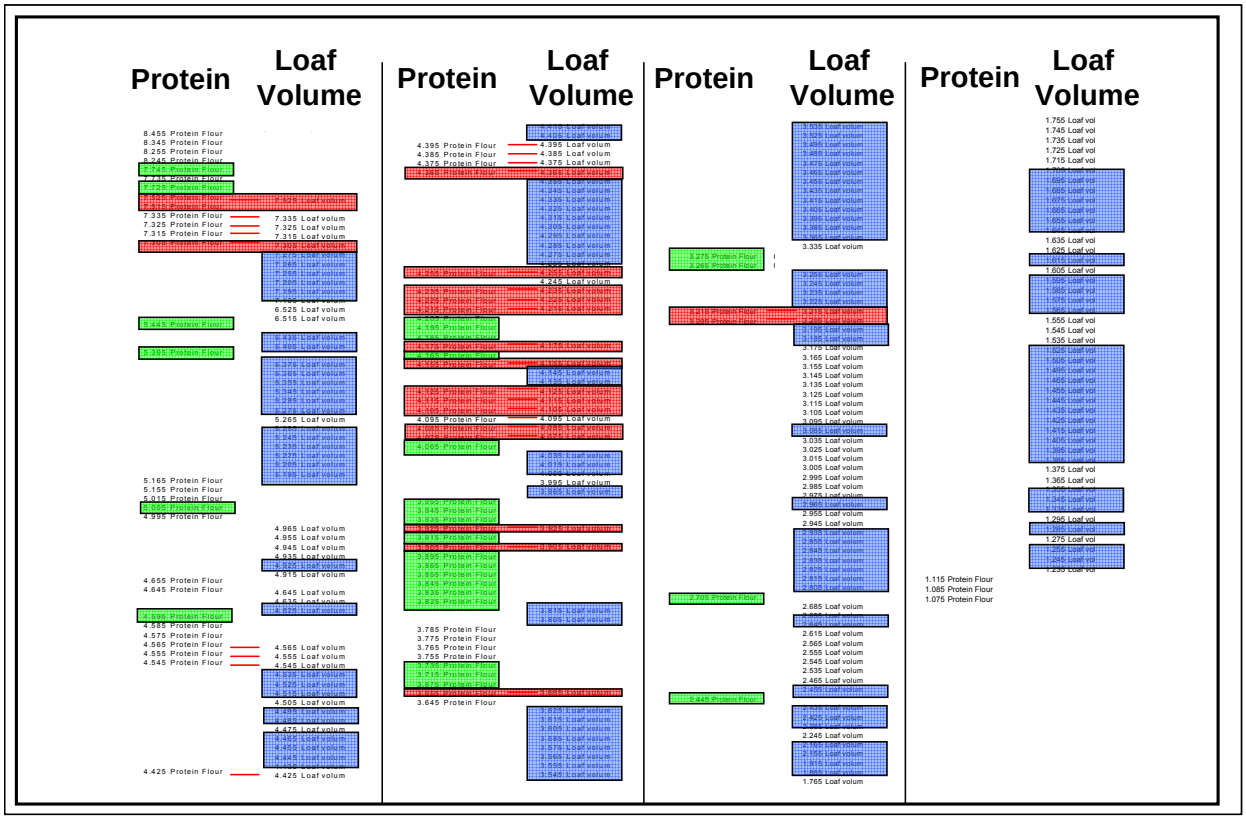
Figure 29. Top ten of the chemical shifts showing the lowest p-value, showing correlations with loaf volume



Obviously, the chemical shifts displaying these correlations may simply arise from a correlation to protein content as loaf volume is known, in some cases, to correlate with protein content. In order to explore this hypothesis, those chemical shifts showing a correlation with protein content were determined. An aligned table, containing all the chemical shifts with p-values below 0.05, was then created to explore the cross-over of loaf volume and protein level correlating chemical shifts (Figure 30).

As can clearly be seen, many of the chemical shifts showing a significant difference did not correlate to either protein content or loaf volume. Other shifts were seen to correlate to loaf volume and protein content. Interestingly some of the shifts correlated only with protein content, with the most significant of these being glycine betaine which showed a strong positive correlation to protein content across all the samples. Of the chemical shifts showing correlations to loaf volume, independently of protein content, two groups were evident. The first group showed a strong negative correlation to loaf volume and consisted of all the peaks corresponding to maltose.

Figure 30. Exploration of correlating chemical shifts to determine cross-over correlations between loaf volume and protein content



Chemical Shifts Correlating with Protein Content and Loaf Volume

Chemical Shifts Correlating with Protein Content Only

Chemical Shifts Correlating with Loaf Volume Only

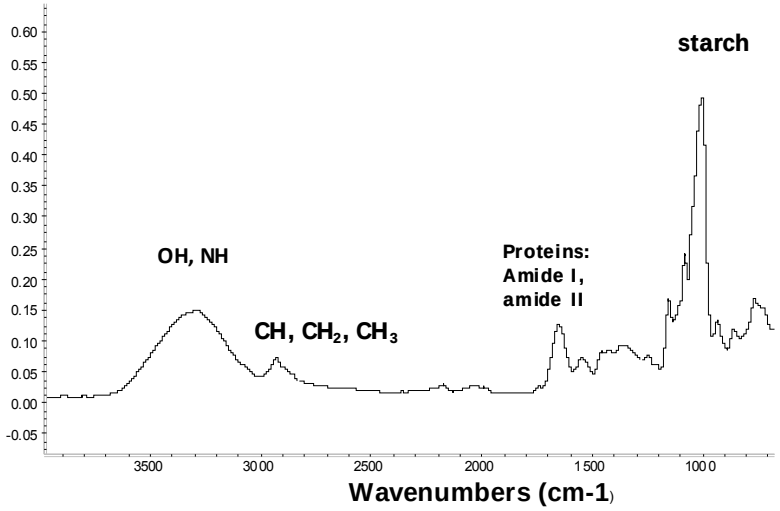
The other group contained chemical shifts showing a positive correlation to loaf volume. These chemical shifts were compared to standards run under the same conditions. It was confirmed that although the chemical shifts resided in the carbohydrate region of the spectrum, they did not correspond to any of the common free sugars contained within the library. The nearest match was to talose but it is more likely that the peaks actually correspond to a conjugated carbohydrate compound not present in the library.

3.8 Fourier transform infrared spectroscopy

Handling of dough samples is recognised to cause changes in gluten structure and as a result, the method of presentation for FT-IR analysis is critical in ensuring reproducible results. To avoid the handling problem and to average as much of the biological variation as possible, it was originally decided to examine the flours derived

from milling without further treatment. A typical spectrum from a dry flour run as received is shown in Figure 31.

Figure 31. An infrared spectrum of Charger flour

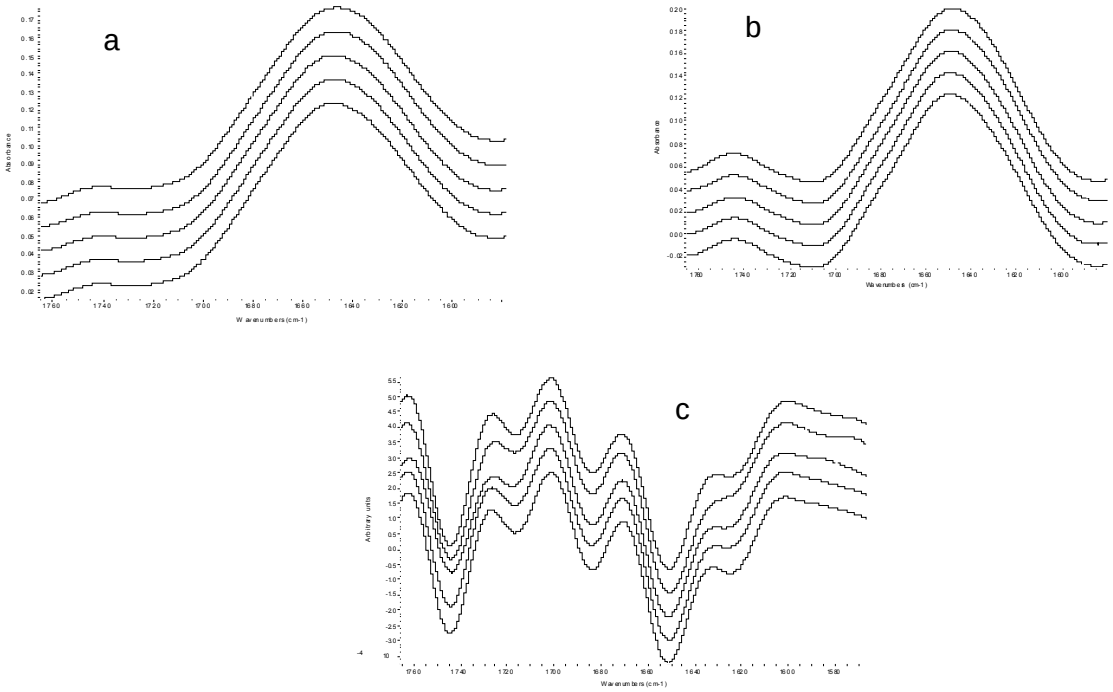


The region of interest is the amide region which is especially sensitive to conformational changes in the protein structure. Details of this region are shown in Figure 32a. In order to make the underlying structure in the spectra clearer, Fourier self deconvoluted and second derivative spectra were obtained (Figures 32b and 32c). In the second derivative spectra the minima represent points where the peaks underlying the envelope have maxima. It is clear from the data that the reproducibility of these spectra is good. However, comparative studies of flour samples from different parents showed little variation.

The gluten in dough is in a hydrated rather than a dry state and it was decided to examine the spectra of hydrated material. As the degree of hydration of the gluten depends on the dough mix and the water absorption properties of the dough and it is known that water content strongly affects the conformation of gluten (Georget and Belton, 2006) it was necessary to choose a suitable hydration level for the experiment.

The only level which could be considered equivalent for all samples was one of maximum hydration in which the flour was exposed to excess water. Under these conditions the proteins would be in their most fully relaxed state.

Figure 32. Five replicate spectra of the amide I region of Charger flour: a, as obtained; b, after Fourier self deconvolution; c, second derivative spectra



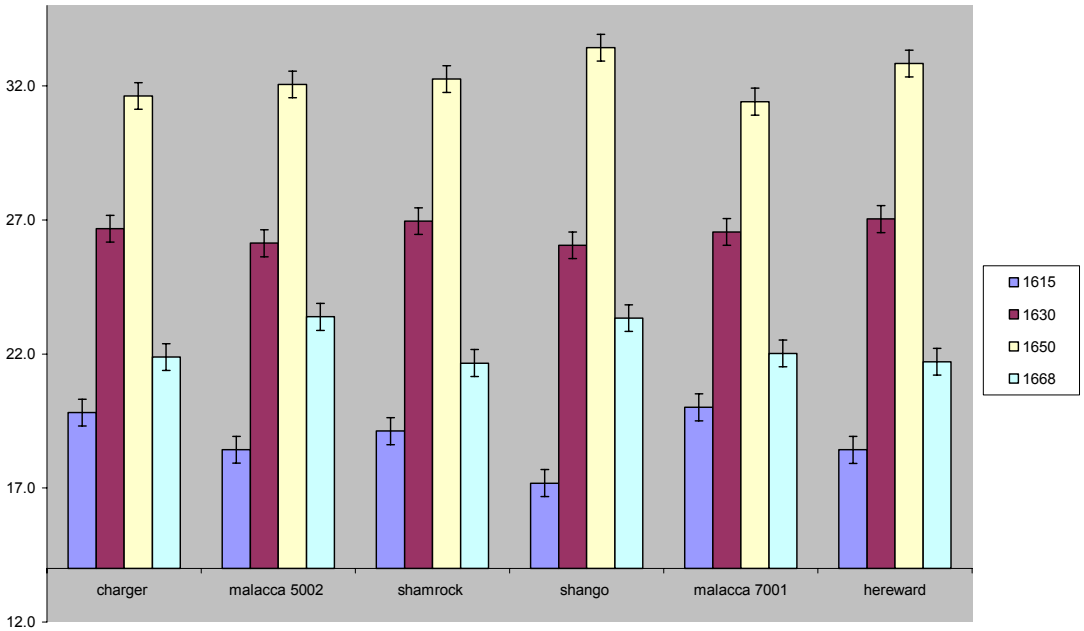
Data obtained under these conditions require the additional spectral manipulation of the subtraction of a water signal that interferes with the Amide I band. This inevitably introduces an increased potential error in the data.

In order to characterise the data, four frequencies were selected that were found to vary significantly in other studies of gluten (Georget and Belton, 2006; Wellner et al., 2006). These are listed in Table 22. The relative intensities of the fully hydrated and dry flours were expressed as a percentage of the sum of the intensities at the four frequencies. This allows direct comparison of the relative intensities in all samples (Figure 33). It is clear from the data that the differences between the samples for each of the frequencies are very close to the standard errors of measurement. One useful way to summarise the data is to take the mean value of the normalised intensity for all samples at a particular frequency and to subtract it from the particular values at that frequency for specific samples. In this way a pattern of deviations from the mean can be established for comparison. Results are shown in Figure 34.

Table 22. Assignment of selected frequencies for gluten dry and hydrated in H_2O. Values in brackets are for the samples hydrated with D_2O

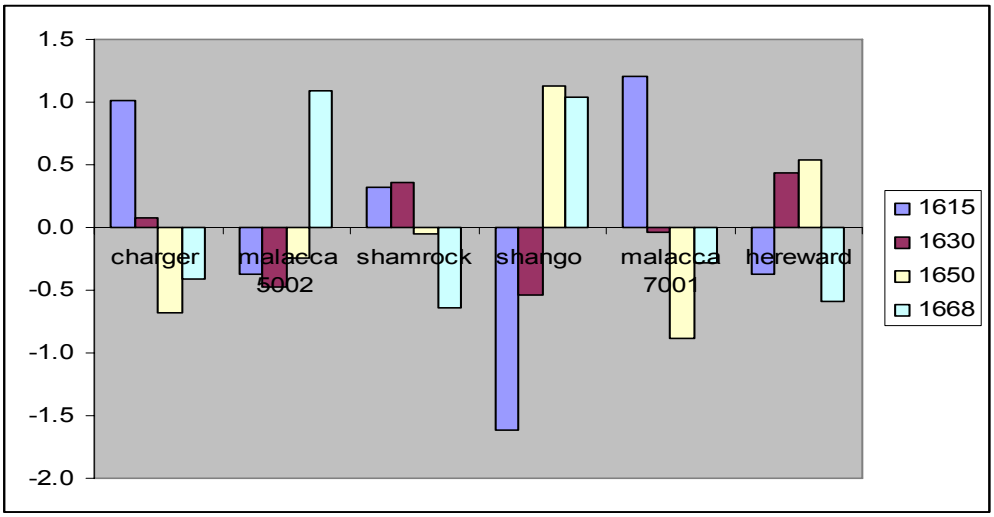
Frequency (cm^{-1})	Assignment
1615(1614)	Strongly hydrogen-bonded *beta* sheets, *beta* edges, extended hydrated chains, some possible contribution from glutamine side chains, intermolecular *beta* sheets
1630(1627)	Antiparallel *beta* sheets, more weakly hydrogen-bonded *beta* sheets
1650(1649)	Random coils and *alpha* helices
1668(1672)	*Beta* turns

Figure 33. Relative intensities at four frequencies for fully hydrated flours. Error bars represent standard deviations of 5 replicates



It is interesting to note in Figure 34 that the two Malacca samples show strong differences in pattern: Malacca from the M×C population (line 5002) shows a strong positive deviation for *beta* turn (1668cm^{-1}) whereas Malacca from the H×M population (line 7001) shows a slight negative deviation. The reverse is true for the 1615 *beta* sheet region.

Figure 34. Deviations of intensities at selected frequencies from the mean of the samples hydrated with H_2O

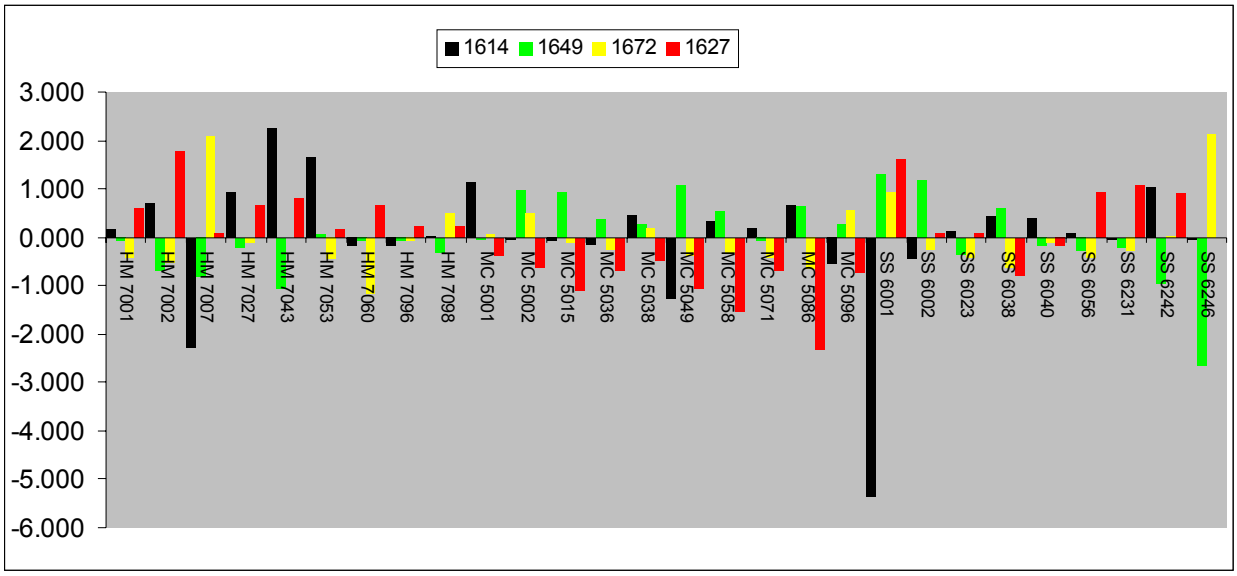


Whilst these differences are suggestive they are not conclusive as they are very close to the limits of error. A second study was undertaken with a larger group of samples using a new spectrometer with considerably improved signal to noise ratios and hydrating with D_2O instead of H_2O. This avoids the need to subtract a water signal, although it does cause small shifts in the Amide I band and effectively removes the Amide II band. Using the same approach as previously, a similar plot to that in Figure 34 was obtained and is shown in Figure 35. The wavelengths used are slightly different to those used with the H_2O treated samples due to small shifts caused by the use of D_2O. However, the assignments are the same as in Table 23.

The data suggest that the series MC 5001 to MC 5096 show negative deviations for the *beta* sheet region (1627cm^{-1}) and positive deviations for the *alpha* helical (1649cm^{-1}) region, whereas the HM series tend to show the reverse. It is notable that

the two Malacca samples still show distinct differences, indicating a role for the agronomic as well as the genetic environment.

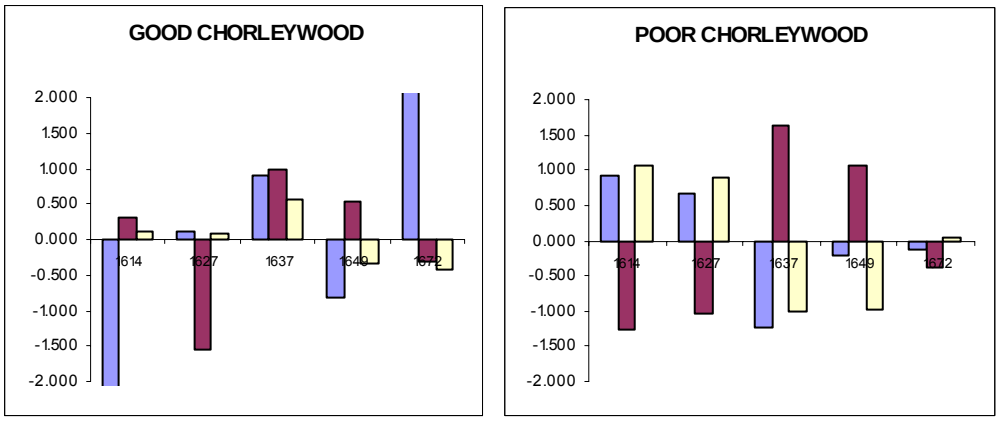
Figure 35. Deviations of intensities at selected frequencies from the mean of the samples hydrated with D_2O



If the results correlate with factors related to baking quality then it would be expected that the deviations of samples with similar baking quality would be similar. In Figure 36 a plot of deviations of samples showing good and poor CBP characteristics is shown. It is clear that the deviations show no sensible correlation with baking quality. Similar results were obtained for wide range of tests (data in Appendix 3) thus indicating that FTIR as used here is not a good indicator of baking quality.

The molecular origins of the spectral variations are difficult to deduce as the distribution of the various proteins in the samples will strongly affect the spectra. *Alpha* and *beta* gliadins typically contain both *alpha* helical and *beta* sheet structures where HMWG subunits and *omega* gliadins have predominance of *beta* sheet and turn structures. It is also likely that the details of the variations in sequence among the subgroups of the major components will also have an effect.

Figure 36. Deviations of samples showing good and poor quality characteristics for the CBP white bread



Further work would, therefore, be required on the detailed analysis of the composition of the glutens used. However a caveat to any future conclusions must be that the differences between intensities in even the best data are very close to the error limits and careful statistical analysis may be required to determine the validity of any differences observed.

4. Summary of results

- Three doubled haploid populations were developed and grown over two harvest years, 2005 and 2006.
- Wheat from both harvests represents a strong dataset showing evidence of transgressive segregation for key parameters.
- Significant variation in processing performance was seen over both years indicating that lines had been well-selected during population development.
- Differences in performance between harvest years was observed, particularly for the Shango×Shamrock population.
- Objective methods developed and used for product assessment were consistent in response and gave a solid platform for subsequent analysis.
- A number of new, statistically significant QTL for milling and baking performance were identified.
- These QTL will form the basis of new breeding initiatives to further improve the quality of UK wheat.
- The relevance of the approach undertaken was underlined using data from routine harvest quality testing which showed that standard methods alone were not wholly effective in predicting functionality.
- Metabolomics analysis showed a number of consistent points in NMR spectra which were related to bread baking performance.
- The relationship between NMR data and loaf volume was not reliant on basic protein content effects and further work will be required to investigate this following completion of the current project.
- Some differences in FT-IR spectral response for the different populations and growing locations were observed.
- FT-IR had previously been effective at assessing wheat protein changes during processing but the technique was less acute in assessing differences between individual samples.

5. Conclusions

This project was unique in scope and depth of analysis of breadmaking quality in wheat. No other groups, worldwide, have been able to put together a project which brings together the combinations of processors, breeders, cereal technologists and geneticists to produce the detailed information on the genetic control of the different breadmaking processes, and the stability of QTL across products and seasons.

From the high quality data produced by the end users, new and novel QTL were discovered for the different traits underlying good breadmaking quality. 194 QTL (at the 5% significance level) were discovered across products and years in the Malacca × Charger population, 179 in Hereward × Malacca population and 233 in the Shango × Shamrock population. It is not surprising that most allelic differences were detected in the Shango × Shamrock cross, since this was the most diverse examined. Gratifyingly, QTL for the same traits were observed across years and where they could be matched, across products, indicating that these data are robust and repeatable.

Many of the QTL discovered can now be targets for marker-assisted selection by the breeders to put together new combinations of alleles which should give better and more consistent quality in UK winter wheat genetic backgrounds. Indeed, the presence of transgressive segregation in the mapping populations for many of the traits measured indicates that genotypes better than the parents can be produced by directed breeding.

These data have moved a long way to understanding the 'Hereward Conundrum' – that is why Hereward is such a good breadmaking quality wheat despite not having excellent high-molecular weight sub-units. It appears that the quality of Hereward is due to novel combinations of several alleles carried at loci not previously identified as being involved in quality variation. This also indicates that these new combinations can be repeated by directed breeding.

6. Acknowledgements

Grateful thanks are due to the individual representatives, past and present, of the following organisations who formed the consortium and played a key role throughout the project in supporting and advising the project leaders at CCFRA as well as undertaking practical work within the project: ADM Milling Ltd, Allied Technical Centre Ltd, Defra, the Federation of Bakers, Heygates Ltd, HGCA, John Innes Centre, National Association of British & Irish Millers (**nabim**), Nickerson (UK) Ltd, RAGT Seeds Ltd, RHM Technology, Rothamsted Research, Smiths Flour Mills, Syngenta Seeds Ltd and the University of East Anglia.

The significant contributions of the previous project leaders at CCFRA, Dr Sybil Barton and Mrs Sue Salmon are gratefully acknowledged.

Thanks are due to HGCA and CEL for the use of the data from samples undergoing RL trials.

7. References

Alava JM, Millar SJ and Salmon SE 2001. The determination of wheat breadmaking performance and bread dough mixing time by NIR spectroscopy for high speed mixers. *Journal of Cereal Science* **33**: 71-81.

Almutawah A, Barker SA and Belton PS 2007. Hydration of gluten: A dielectric, calorimetric, and Fourier transform infrared study. *Biomacromolecules* **8** (5): 1601-1606.

Anderson CA 2003. Characterising wheat flour protein quality from Reomixer traces. HGCA Project Report No. 324.

Andersson R, Hämäläinen M and Åman P 1994. Predictive modelling of the bread-making performance and dough properties of wheat. *Journal of Cereal Science* **20**: 129-138.

Bassam BJ, Caetano-Anollés G and Gresshoff PM 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* **196**: 80-83.

CCFRA 2002. Guideline No. 3 Manual of Methods for Wheat and Flour Testing 3rd Edition.

Georget D M R and Belton P S 2006. Effects of temperature and water content on the secondary structure of wheat gluten studied by FTIR spectroscopy. *Biomacromolecules* **7** (2): 469-475.

Groos C, Bervas, E, Chanliaud and Charmet, G 2007. Genetic analysis of bread-making quality scores in bread wheat using a recombinant inbred line population. *Theoretical and Applied Genetics* **115**: 313-323.

Jaccoud D, Peng K, Feinstein D and Kilian A 2001. Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* **29**:No.4 e25.

Kuchel H, Langridge, P, Mosionek, L, Williams, K and Jeffries, SP 2006. The genetic control of milling yield, dough rheology and baking quality of wheat. *Theoretical and Applied Genetics* **112**:1487-1495.

Law CN, Bhandari DG, Salmon SE, Greenwell PW, Foot IM, Cauvain SP, Sayers EJ and Worland AJ 2005. Novel genes on chromosome 3A influencing breadmaking quality in wheat, including a new gene for loaf volume, *Lvl 1*. *Journal of Cereal Science* **41**:317-326.

Millar SJ 2003. The development of near infrared (NIR) spectroscopy calibrations for the prediction of wheat and flour quality. HGCA Project Report No. 310.

nabim 2007. UK Flour Milling Industry 2007. The National Association of British and Irish Millers.

Payne P 1987. Genetics of wheat storage proteins and the effect of the allelic variation on bread-making quality. *Annual Review of Plant Physiology* **38**:141-153.

Payne P I, Nightingale M A, Krattiger A F and Holt L M 1987. The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *Journal of the Science of Food and Agriculture* **40**: 51-65.

Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD and Ganal MW 1995. Abundance, variability and chromosome location of microsatellites in wheat. *Molecular and General Genetics* **246**: 327-333.

Rousset M, Brabant P, Kota, R, Dubcovsky J and Dvorak J 2001. Use of recombinant substitution lines for gene mapping and QTL analysis of breadmaking quality in wheat. *Euphytica* **119**: 81-87.

Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y and Ciavarrella M 2007. Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* **154**: 401-408.

Somers DJ, Isaac P and Edwards K 2004. A high-density wheat microsatellite consensus map for bread wheat (*Triticum aestivum L.*). *Theoretical and Applied Genetics* **109**:1105-1114.

Wellner N, Mills E N C, Brownsey G, Wilson R H, Brown N, Freeman J, Halford N G, Shewry P R and Belton P S 2005. Changes in protein secondary structure during gluten deformation studied by dynamic Fourier transform infrared spectroscopy. *Biomacromolecules* **6** (1): 255-261.

Wellner N, Marsh J T, Savage A W J, Halford N G, Shewry P R, Mills E N C and Belton P S 2006. Comparison of repetitive sequences derived from high molecular weight subunits of wheat glutenin, an elastomeric plant protein. *Biomacromolecules* **7** (4): 1096-1103.