March 2014

# Project Report No. 528

## Association genetics of UK elite barley (AGOUEB)

by

W Thomas[1], J Comadran[1], L Ramsay[1], P Shaw[1], D Marshall[1], A Newton[1], D O'Sullivan[2], J Cockram[2], I Mackay[2], R Bayles[2], J White[2], M Kearsey[3], Z Luo[3], M Wang[3], C Tapsell[4], D Harrap[4], P Werner[4], S Klose[5], P Bury[6], J Wroth[6], O Argillier[6], R Habgood[7], M Glew[7], A-M Bochard[7], P Gymer[8], D Vequaud[8], T Christerson[9], B Allvin[10], N Davies[11], R Broadbent[11], J Brosnan[12], T Bringhurst[12], C Booer[13] and R Waugh[1]

[1]SCRI, Invergowrie, Dundee DD2 5DA
[2]NIAB, Huntingdon Road, Cambridge CB3 0LE
[3]University of Birmingham, Edgbaston, Birmingham B15 2TT
[4]KWS UK, 56 Church Street, Thriplow, Nr Royston, Herts SG8 7RE
[5]LS Plant Breeding, North Barn, Manor Farm, Milton Road, Cambridge CB24 9NG
[6]Syngenta Seeds Ltd, Market Stainton, Market Rasen, Lincs LN8 5LJ
[7]Nickerson-Advanta Ltd, Joseph Nickerson Research Centre, Rothwell, Market Rasen, Lincs LN7 6DT
[8]Secobra/Dalgety, Throws Farm, Stebbing, Great Dunmow, Essex CM6 3AQ
[9]SW Seed, Svalöf Weibull AB, SE-268 81 Svalöv, Sweden
[10]Perten Instruments AB, P.O. Box 5101, SE-141 05 Huddinge, Sweden
[11]The Maltsters Association of Great Britain, 31B Castlegate, Newark, Notts NG24 1AZ
[12]The Scotch Whisky Research Institute, The Robertson Trust Building, Research Park North, Riccarton, Edinburgh EH14 4AP
[13]Campden BRI, Lyttel Hall, Nutfield, Surrey RH1 4HY

This is the final report of a 54 month project (RD-2004-3043) which started in October 2005. The work was funded by the Sustainable Arable Link programme and a contract for £161,066 from HGCA.

While the Agriculture and Horticulture Development Board, operating through its HGCA division, seeks to ensure that the information contained within this document is accurate at the time of printing, no warranty is given in respect thereof and, to the maximum extent permitted by law, the Agriculture and Horticulture Development Board accepts no liability for loss, damage or injury howsoever caused (including that caused by negligence) or suffered directly or indirectly in relation to information and opinions contained in or omitted from this document.

Reference herein to trade names and proprietary products without stating that they are protected does not imply that they may be regarded as unprotected and thus free for general use. No endorsement of named products is intended, nor is any criticism implied of other alternative, but unnamed, products.

HGCA is the cereals and oilseeds division of the Agriculture and Horticulture Development Board.

# CONTENTS

# 1. ABSTRACT

Our aim was to DNA fingerprint 1000 barley genotypes using a panel of 3000 molecular markers to characterise the variation that exists amongst UK elite barley varieties and associate variation in marker profiles i.e. DNA fingerprints with differences in performance and morphological characters. Over 500 of these lines had been evaluated in spring and winter barley National and Recommended List trials between 1988 and 2006 and thus an extensive body of performance data (yield, height, disease resistance, quality etc.) already existed for these lines. Additionally we grew a subset of lines representing market successes and failures over our survey period in a series of trials to provide an unambiguous estimate of breeding progress and additional data to improve the prediction of means of varieties that generally were not grown in the same trials.

Multi-variate analysis of the marker data generated by the DNA fingerprinting resulted in three general groupings that represented spring barley, two row winter barley and six row winter barley. Whilst some varieties were genetically (and morphologically) quite similar, e.g. Angora and Melanie, there were still considerable genetic differences between varieties within the three major groupings and therefore plenty of genetic variation for breeders to continue to exploit. This was borne out by the results of the trials carried out within the project, where we clearly demonstrated breeding progress for yield, which appeared to be due to increased grain size, in the winter and spring crop and also for malt extract in the spring crop.

We have also been able to identify associations of individual molecular markers with morphological characters used to establish Distinctness, Uniformity and Stability (DUS) in National List testing. For the first time, we can therefore confirm that DUS characters are controlled by genes on each of barley's seven chromosomes. We have also been able to closely define the chromosomal region harbouring these controlling genes and use the similarity between barley and the fully sequenced genomes of rice and *Brachypodium* to identify potential candidate genes for the characters. The success of this approach has been demonstrated by the cloning of the major gene responsible for the development of anthocyanin pigmentation in various barley tissues. The approach is also proving successful in identifying genes responsible for performance characters such as yield. Whilst this offers the prospect of utilising DNA markers to predict the favourable combinations of alleles as a short-cut to directly measuring performance, a technique known as Marker-Assisted Selection, we still need to determine how all these genes act together to produce enhanced performance.

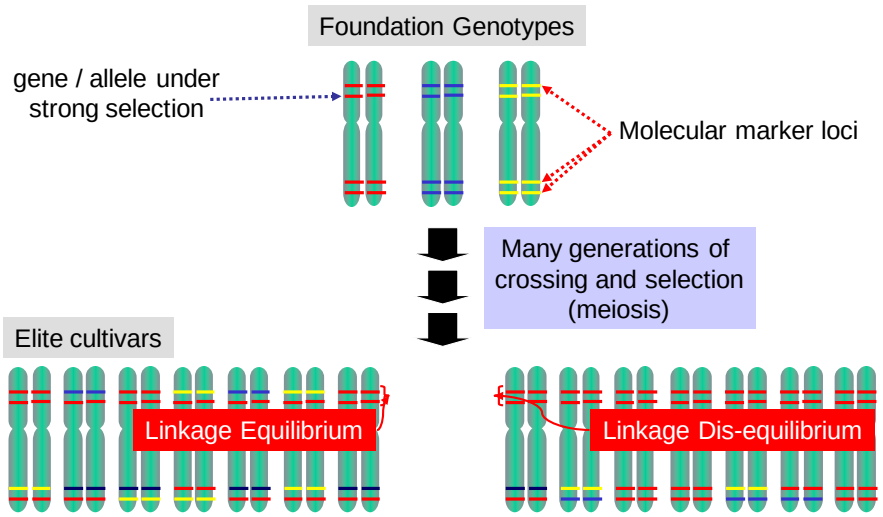
Finally, the extent of the DNA fingerprinting that we have carried out means that future experiments can be based upon the same material that we have studied but could access the existing molecular marker information, removing the need and expense of large-scale genotyping material in the experiment as well as phenotyping.

# 2. SUMMARY

## 2.1. Background

Plant breeding is a lengthy and expensive Research and Development exercise with selection for key performance characters such as yield and quality based upon direct measurements from trials that are generally grown in the later stages of the selection process. The use of molecular markers that are 'genetic signposts' for genetic regions associated with improvements in performance can be applied in the early stages of the selection process (Marker Assisted Selection) and provide an enriched gene pool from which the best of the best lines can be selected by direct measurements. Marker Assisted Selection therefore offers the promise of improved efficiency of plant breeding but, despite years of research, has yet to be realised on a large scale for small-grained cereals. Previous plant research to identify markers to use in Marker Assisted Selection have concentrated on progenies from specific crosses due to their practical ease and genetical and statistical power to detect significant associations of markers with characters, so-called Quantitative Trait Loci (QTL). Crosses were frequently chosen that maximise parental differences in order to facilitate whole genome map construction and the detection of genetic regions affecting characters such as yield and agronomic performance. Whilst this produces significant results, the applicability of such findings in general plant breeding was very limited because the crosses were far from typical of those being made in current breeding programmes. Commercial plant breeding programmes are generally based upon inter-crosses of quite similar lines where the differences are not so marked but nevertheless result in genetic improvements. In contrast, mammalian genetic studies cannot utilise such special populations and have developed approaches based upon assembling a diverse collection of individuals and collecting phenotypic data upon them. The collection is also genotyped with an extensive set of DNA markers, allelic variation at each of which is tested to determine if it is significantly associated with the observed phenotypic differences. This approach is termed association mapping and potentially can be used to assess all the differences that exist within a collection of lines, e.g. all the winter barley lines that have been recommended over the past 40 years. Effectively, the approach relies upon detecting DNA markers that are closely linked to chromosomal segments that are under selection because they contain a gene (or genes) that control important characters such as yield. In such cases, a specific allele of the closely linked marker tends to be in excess in the higher yielding lines and another allele in excess in lower yielding lines. Such a case is termed Linkage Disequilibrium as the expectation is that marker alleles would be equally distributed amongst higher and lower yielding lines if there was no association – Linkage Equilibrium. In plants, the challenge is to generate sufficiently detailed genetic fingerprints in order to detect significant marker/character associations before recombination means that allelic differences at a marker return to equilibrium (Summary Figure 1).

It follows that the power and discrimination of association studies is largely determined by the underlying patterns of linkage disequilibrium (LD) in the chosen population, which is a variable that can be affected by one or more of the following factors: a.) past history of the population, b.) domestication bottle-necks, c.) selection, d.) population admixture, and e.) population structure. In addition, detection will also depend upon aspects of the phenotype, such as heritability, penetrance, size of the effect etc. If, however, LD can be used in association genetics of pools of elite germplasm that barley breeders are working with then the results will be immediately applicable and facilitate more effective selection with downstream benefits to growers and processors as well as end-users. It follows from the previous paragraph that for association genetics analyses to be effective, LD must persist to a greater degree than the average marker density. For instance, barley chromosomes should have a genetic map length of 1200 centimorgans (cM) and a typical survey with 120 Simple Sequence Repeat (SSR) markers means that LD would need to persist beyond 10cM to have a chance of detecting significant association of an SSR with a performance trait. Whilst this is possible in barley, increasing the marker density would not only increase the reliability of detection but also resolve the chromosomal region controlling a performance trait to a much narrower region and thus help break down adverse tight linkages.




**Summary Figure 1.** Derivation of inbred lines (represented by a single pair of chromosomes) with markers either associated with a character under selection (Linkage Disequilibrium, lower RHS) or not associated (Linkage Equilibrium, lower LHS).

## 2.2. Methodology and objectives

We viewed association genetics analyses as the most promising methodology to identify the location of genes controlling important characters for growing and processing barley and thereby identify molecular markers that could be used by barley breeders to select varieties that better meet these two needs. We therefore needed to combine a genetic fingerprinting methodology with high quality phenotypic data.

### 2.2.1. Genetic fingerprinting

Previous studies had shown that most of the variation amongst elite spring barley varieties originated from relatively few 'founder' varieties. This means that there is a relatively small number of alleles at any one gene and very limited variation around it. A consequence of this is that different varieties tend to share the same pattern of allelic differences, or haplotype, in any one chromosomal segment. Preliminary studies showed that linkage disequilibrium can persist for distances of up to several cM. Association genetics studies with markers spaced on average 1cM apart should therefore permit the identification of chromosomal regions affecting most characters. Prior to the project, such a scan was limited by the lack of a high throughput and low-cost genotyping assay of over 1000 markers that could be applied to a large number (1000) genotypes at the same time.

Through international collaboration, we were able to identify a large number of Single Nucleotide Polymorphisms (SNPs) in individual barley genes, which we then used to develop a high-throughput SNP genotyping platform for barley, using the Illumina Goldengate assay. Three pilot-phase 1536 SNP GoldenGate Oligo Pooled Arrays (OPAs) assays referred to as POPA1, POPA2 and POPA3 were developed and two 1536 SNP production scale OPAs, referred to as BOPA1 and BOPA2, were developed from SNPs tested on the pilot OPAs that then formed the basis of the genotyping undertaken in the project. BOPA1 consisted of single SNPs within 1536 separate genes and BOPA2 consisted of single SNPs within additional genes, additional SNPs within some of the genes on BOPA1 and specific SNPs designed to sample the known variants at key developmental loci such as Ppd-H1 and Vrn-H2, as well as the *mlo* mildew resistance locus.

### 2.2.2. Phenotypic data

Barley breeders have continued to improve the yield and agronomic potential of barley as witnessed by the increases offered by new additions to the HGCA/CEL Recommended Lists (RL) of varieties for growing in the UK. The phenotypic data underlying these improvements is very extensive with over 20 fungicide treated trials for spring and winter barley each year. There are generally 7-8 treated trials each year for the stage before RL, National List (NL) trials, so that the mean performance of a line that has completed two years of NL will be based upon at least 14 site

x season combinations. Most research studies of yield and agronomic performance are based upon 4-8 site x season combinations so the combination of data from National and Recommended List trials represents a major improvement in the range of environments sampled. Examination of the HGCA Recommended Lists show that the standard errors associated with the means for characters such as yield are very small compared to those achieved in normal QTL studies and the extra precision of these data increased our chances of identifying genetic regions controlling them.

No one genotype was in trials throughout the period of study, so it was necessary for some varieties to be grown over overlapping periods of at least five years to provide reliable predictions of genotypic means for material grown in different seasons. This was generally provided by the designated control varieties for National List trialling, augmented by other successful varieties that remained on the Recommended List for a long period. To improve the accuracy of these predictions, we identified a number of winter and spring barley lines for growing in trials at a number of sites for the first three harvest years of the project. These included some that had been market successes, some failures, some that failed to gain a recommendation and some key progenitors. As the micro-malting protocol had changed during our survey period, these trials also provided material for end user quality analyses through collaboration with the Maltsters Association of Great Britain, the Scotch Whisky Research Institute and Campden Brewing Research International.

### 2.2.3. Objectives

We therefore set the following seven objectives for the project:

1. To genotype 1000 elite barley lines with 3000 SNP marker loci developed from 1000-1500 genes (i.e. 2-3 SNPs/gene) to give the 1-2cM resolution necessary for association genetics studies in barley.
2. To assemble and analyse phenotypic data for economically important traits from extant official trial and breeders' data.
3. To generate novel phenotypic data through additional field, glasshouse and controlled environment studies. In particular these will include a detailed characterisation of disease resistance.
4. To analyse the genotypic data to determine the level and extent of linkage disequilibrium across the genome, to predict haplotype structures and to test for various factors such as population structure that shape the observed linkage disequilibrium patterns.
5. To investigate the association of the genotypic and phenotypic data to determine the genetic control of qualitative and quantitative traits. The latter will include highly heritable traits such as heading date (flowering) for which several candidate genes/genomic regions are known and other biologically and agronomically important traits such as yield and yield components that have proved less tractable to date using traditional linkage studies.

6. Development of theoretical approaches for modelling linkage and linkage disequilibrium between polymorphic markers and quantitative trait loci in the populations with and without genealogical information, and development of data-specific statistical methods and computational tools for analyzing the data generated in the project.
7. To develop a GERMINATE based web accessible public/private database that will store the genotypic and phenotypic data and the results of association genetic analysis.

## 2.3. Results

**Objective 1 – Collection of genotypic data.** We genotyped 1,000 spring and winter barley lines with 1536 BOPA1 and 1536 BOPA2 SNP markers plus another 24 with BOPA1 only. These genotypes comprise 547 in a publicly available set (the AGOUEB Public Set of RL and NL lines), 117 progenitor and disease differential lines and 360 breeders 'private' lines. Results from BOPA1 were generally better than those from BOPA2 in terms of the success rate of the markers. In addition, we identified some problems in the consistency of the genotypic data sets from BOPA1 and BOPA2, indicating some seed and/or DNA and/or data handling problems, resulting in the elimination of BOPA2 genotypes for a number of lines.

**Objective 2 – Phenotypic data.** We assembled data from the means of 849 fungicide treated trials grown over harvest years 1988-2006 as part of the National List and Recommended List Trial series for 282 spring and 325 winter lines for a number of variates.

**Objective 3 – Collection of new phenotypic data.** We selected 67 spring and 63 winter barley varieties to represent key varieties and important parents from the 1960s to the 2000s. These were grown in trials by SCRI, NIAB and the breeder partners to provide eight spring and seven winter barley sites for each of the harvest years 2006-2008.

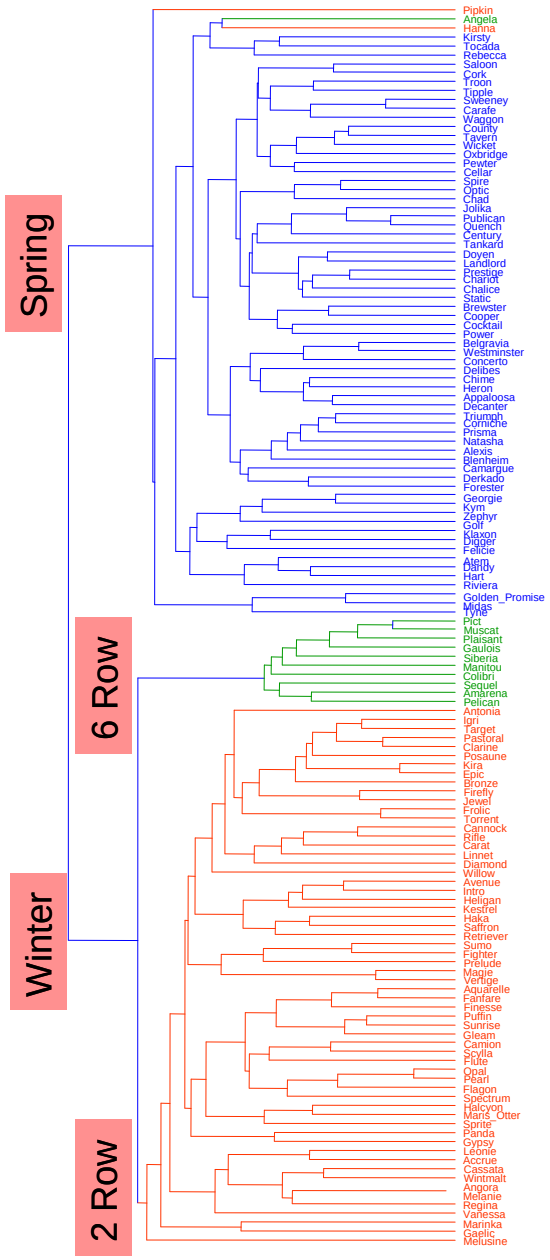
Apart from agronomic yield, the data coverage in the National and Recommended List Trials was surprisingly poor for other variates and we did not consider any variates for which there were less than 1000 data points. This left us with 44 and 39 variates from spring and winter barley trials respectively from which we derived overall means for each genotype in a mixed model analysis where we also accounted for differences between years and sites and the various interactions between the three main effects. By including the trials grown for Objective 3 as well as the data assembled under Objective 2 in the same analysis, we obtained means for up to 295 spring and 333 winter lines. We were able to obtain seed of 547 of these 628 lines and this was therefore the maximum number of lines for association analysis of the phenotypic data. These lines had all been in multi-site trials for at least two years and the data from them was shared amongst all project partners with the aim of making the data public at some time after the end of the project. We therefore called these lines the AGOUEB Public Set whereas data on lines contributed by

individual breeders remained private to the contributing company and was therefore called Breeder X's private set. The average heritability of the spring traits was 35.6% and 44.7% for the winters with a minimum of 5.6% for Germinative Energy in 4ml of water and a maximum of 83.4% for grain width:length ratio, both found for winter barley.

As part of the National Listing process, each line submitted has to be shown as distinct from any other line on the National List, that seed from the line are all the same (uniform) or that there are fewer than a defined limit of off-types, and that seed from the line breed true and thus are stable. This is established by assessing each submission over a period of two years for a range of discrete morphological characters in Distinctness, Uniformity and Stability (DUS) tests. As with the traits collected from the trials, coverage of these traits was not only highly variable but also subject to changes in the scoring for the genotypes surveyed in the project. After unifying scoring scales where necessary and possible, we had 33 DUS traits with scores for at least 200 of the genotypes under study, which we considered to be the minimum set for association analyses. Whilst eight of these were characters that fell into just two classes, e.g. two/six row ears, fourteen fell into nine different classes on a graduated scale, e.g. absence of wax to very strong wax deposition on the ear. One of the traits, disposition of the lodicules, was virtually monomorphic with only five variant phenotypes and was therefore excluded from analyses as the allele frequency was too low.

**Objective 4 – Analysis of the population structure.** Multi-variate analysis of the genotypic data collected upon the AGOUEB Public set revealed up to ten groups. However most of the differences between the genotypes were described by the three main developmental classes that they can be categorised by: spring two row, winter two row, and winter six row.

The three main categories can be seen in the dendrogram derived from BOPA1 genotyping shown in Summary Figure 2 where the 138 varieties from the AGOUEB Public Set that had also been placed upon the UK barley Recommended Lists are displayed. The 65 spring varieties are clearly separated from the 73 winter varieties and, within the latter, the two rows are clearly separated from the six rows.




**Summary Figure 2.** Dendrogram produced by BOPA1 genotyping of 138 barley varieties recommended between 1965 and 2006. Spring varieties are in blue, winter two row in red and winter six row in green.

The two most similar lines were the winter barley varieties Melanie and Angora, which were two sister selections from the same cross by the same breeder. Melanie and Angora were morphologically indistinguishable and could only be separated by examining their hordein strorage protein banding patterns. These two varieties differed at 24 of over 2800 loci where we had marker data for both, i.e. the two lines are over 99% similar. More importantly, we found that these differences were predominantly located on the short arm of barley chromosome 1, the distal portion of the long arm of 5 and the central portion of 6. The region of chromosome 1 is where genes controlling variation in the hordein storage proteins are located and thus the SNP differences in the region are consistent with hordein storage protein differences. The clear genetic differences detected by the SNP markers are in two genomic regions that are not covered by the

DUS markers that we have detected (Summary Figure 3). Whilst this difference seems small, we have sampled probably less than 10% of all barley genes yet over 16 million different genetic combinations can be produced from the 24 segregating loci that we have detected.

The next two most similar lines were the winter varieties Pearl and Opal, which again were sister lines from the same cross by the same breeder. The most similar spring varieties were Publican and Quench, which were selections by the same breeder from a cross made between the same parents but reciprocally. These two varieties differed at over 140 out of over 2800 marker loci where we have genotypic data for both, i.e. they are 95% similar.

Whilst this is a representation of the picture seen in the whole of the AGOUEB Public set, there are three winter varieties that are mis-classified on Summary Figure 2. One is Pipkin which, whilst classified as a winter variety, is a selection from a cross between Maris Otter and Sergeant and thus is likely to have a high proportion of spring alleles. The other two varieties are the six row Angela and the two row Hanna. Both these are varieties with names that have been used previously and closer inspection of the genebank accessions used to source the plant for genotyping revealed that we had accessed the earlier versions of both lines, which were spring two row types. This illustrates not only the problem in being careful to source the right accessions from gene-banks but also highlights the need to ensure that unique identifiers, such as breeders codes are stored in genebanks for all lines entered into National List trials.

The extent of Linkage Disequilibrium was variable across the genome. In centromeric regions, where recombination in barley is known to be limited, Linkage Disequilibrium persisted for over 10cM in some cases, e.g. chromosome 3H. Elsewhere in the genome, Linkage Disequilibrium decayed more rapidly but generally persisted for 2cM so that detection of marker trait associations with our marker density was truly feasible.

### 2.3.1. Associations with DUS characters

We detected significant associations of 41 markers with 19 of the 33 DUS traits. Most of the DUS traits where we found significant associations were associated with just one genomic location (Summary Table 1). The notable exceptions were Winter vs Spring growth habit (Season) and number of rows on the ear (Rows), where we detected 10 and 7 significant locations respectively. Whilst we correctly identified the locations of the two major loci known to affect ear row type on chromosomes 2H and 4H, some of the other predicted locations were surprising. Similarly, we correctly predicted the location of two of the seasonal growth habit loci, VrnH1 and VrnH2 on chromosomes 5H and 4H respectively but the finding of five additional associations in the region of VrnH1 was surprising.

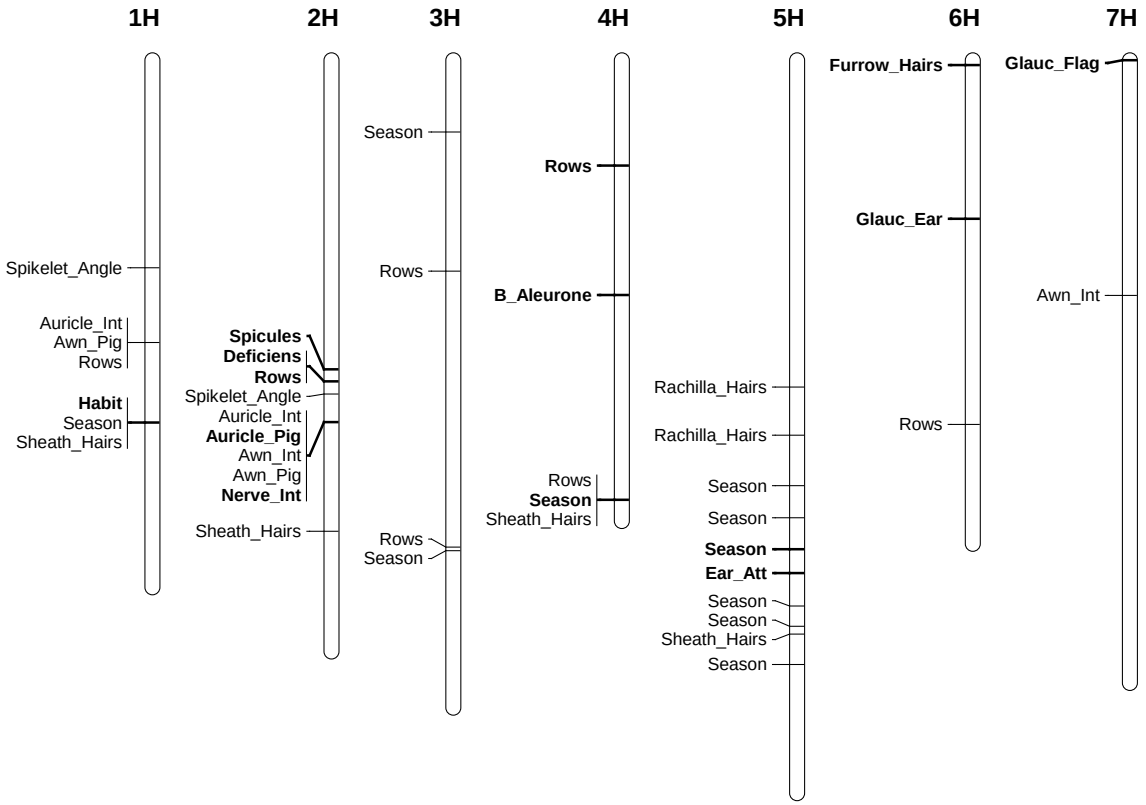
**Summary Table 1.** Number of significant associations detected for 19 morphological characters used to establish Distinctness, Uniformity, and Stability (DUS) in barley

| **Character (Abbreviation)** | **Number of markers significantly associated with a character** |
|---|---|
| Intensity of Auricle Pigmentation (Auricle_Int) | 2 |
| Auricle Pigmentation (Auricle_Pig) | 1 |
| Intensity of Awn Pigmentation (Awn_Int) | 2 |
| Awn Pigmentation (Awn_Pig) | 2 |
| Blue/white Aleurone Colour (B_Aleurone) | 1 |
| Absence/Vestigial Lateral Florets (Deficiens) | 1 |
| Ear Attitude (Ear_Att) | 1 |
| Ventral Furrow Hairs (Furrow_Hairs) | 1 |
| Ear Glaucosity (Glauc_Ear) | 1 |
| Flag Leaf Glaucosity (Glauc_Flag) | 1 |
| Plant Habit (Habit) | 1 |
| Intensity of Lemma Nerve Pigmentation (Nerve_Int) | 1 |
| Length of Rachilla Hairs (Rachilla_Hairs) | 2 |
| 2/6 Row Ears (Rows) | 7 |
| Winter or Spring Seasonal Habit (Season) | 10 |
| Presence/Absence of Leaf Sheath Hairs (Sheath_Hairs) | 4 |
| Spicules on Inner Lateral Nerves of Grain (Spicules) | 1 |
| Spikelet Angle | 1 |
| Grand Total | 41 |

The distribution of the loci controlling the DUS traits is shown in Summary Figure 3. We clearly identified a major locus affecting pigmentation on chromosome 2H and used this information to identify the causal gene and the polymorphism responsible for non-production of anthocyanin. Whilst this appeared a reasonable sampling of the barley genome and many of the loci identified were in plausible regions, e.g. Leaf Sheath Hairs and Blue Aleurone on chromosome 4H, there were also some surprising associations. For instance, the association of Rows and Season on the long arm of chromosome 3H was in the region of the *sdw1* dwarfing gene, which is only found in spring two-rowed types in UK elite germplasm. It seems unlikely that this region of the genome affects these characters and the finding may be due to the assumption of additivity for characters known to be subject to gene interactions (epistatis).

Nonetheless, we have identified major loci controlling many DUS characters and this knowledge will help barley breeders develop new varieties that are distinct from each other. There are also large regions of the barley genome that are not involved in the genetic control of DUS characters

and thus allow barley breeders unconstrained opportunities for recombining to improve phenotypic performance.

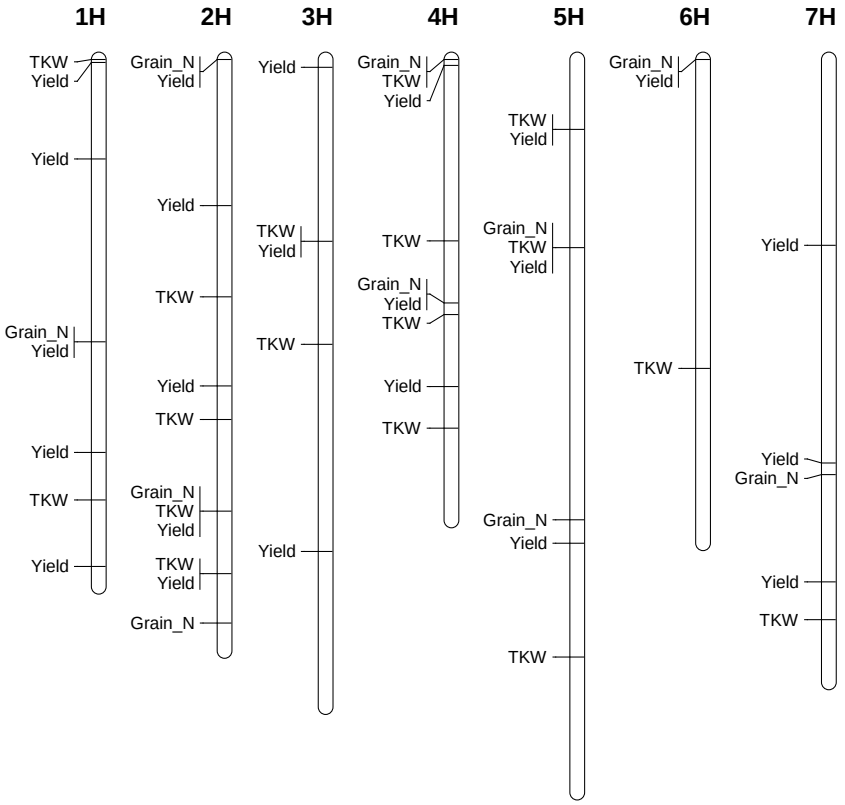



**Summary Figure 3.** Chromosomal locations of marker associations with 19 DUS characters. Characters in bold indicate locations where just one marker is associated with a trait or where a location corresponds to the map position of the morphological character.

### 2.3.2. Associations with VCU characters

We used a kinship matrix constructed from the BOPA1 genotypes to correct for population sub-structure in the spring genotypes and then conducted single marker genome scans to identify significant associations with VCU phenotypic data. Here, we have summarised the findings from studies of the BLUPs derived from an overall analysis of the NL and RL databases (see above). Such an analysis will detect overall mean effects for each locus but, because of the unbalanced nature of the data-set, we cannot detect QTL x Environment interactions, which are likely to be important for some characters. Nevertheless, we detected at least one significant marker/phenotype association for the characters that we studied and the genomic locations of markers associated with Yield, Thousand Kernel Weight and Grain Nitrogen are shown in Summary Figure 4. Most of these associations can be found amongst the large body of published QTL locations from pair crosses and are therefore validated by these reports. Crucially, we have identified all these associations from studying existing data for just one population, whereas the published reports have been accumulated from a large number of separate studies. Thus, our approach has not only provided valuable data but also been extremely cost-effective.

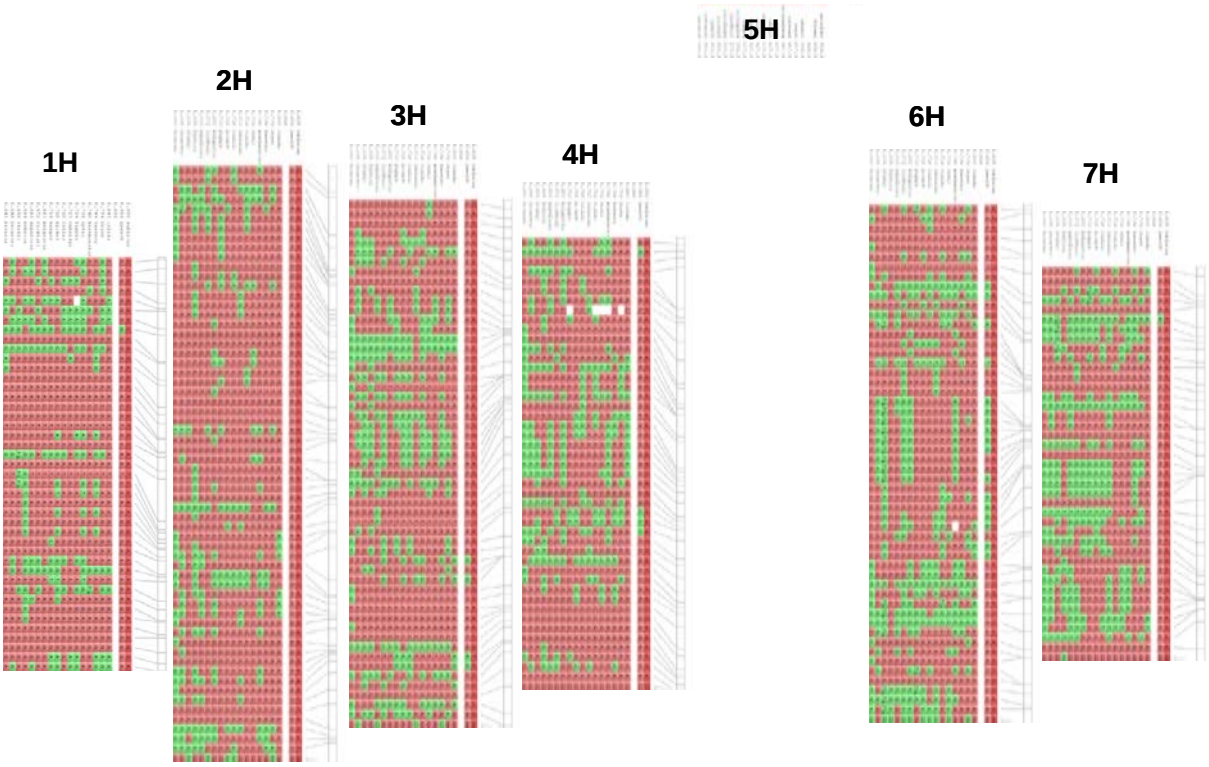
Summary Figure 4 also shows that many of the associations are co-located, reflecting the correlations between characters. For instance, a marker towards the centromeric region of chromosome 5H is associated with all three characters such that an increase in thousand grain weight results not only in an increase in yield but also a decrease in Grain Nitrogen content. Curated BOPA2 data were not available in time to conduct analyses to report here but a fuller analysis of VCU data, especially BLUPs obtained by including the data generated from the AGOUEB trial series, using BOPA1 and BOPA2 markers will be carried out and the scans lodged on the AGOUEB web site (www.agoueb.org).




**Summary Figure 4.** Chromosomal location of markers significantly associated with yield, thousand kernel weight (TKW) and grain nitrogen content (Grain_N) for more than 200 spring barley lines grown in UK National and Recommended List treated trials 1988 - 2006.

**Objective 6 – Alternative analytical methods for association genetics.** Existing and novel approaches to account for population structure were tested and applied to the data generated from the project. Whilst some of the novel approaches appeared promising alternatives, we decided to adopt a standard approach for the analyses of associations with DUS and VCU characters. This was based upon use of a kinship matrix to account for population structure.

**Objective 7 – Develop a database for data storage and retrieval within the project.** We used a GERMINATE database (http://bioinf.hutton.ac.uk/germinate) to store all the data associated with the project as it is designed to hold a range of passport type data to classify entries and a range of additional data types including genotypic data, phenotypic data, and pedigrees. We have also developed a range of associated tools to visualise the data. For instance, we have developed FLAPJACK (http://bioinf.hutton.ac.uk/flapjack) to visualise the matrix of genotypes and markers. Summary Figure 5 has been produced from a FLAPJACK representation of the HGCA/CEL 2008 spring barley Recommended List to highlight the differences between the 19 varieties with Publican and Quench being the two rightmost varieties for each chromosome. Quench is a selection from a cross between the same parents as Publican and differs at 34 of a subset of 384 of the 1536 BOPA1marker loci. The question is whether any of these 34 differences are associated with the fact that Publican is an accepted distilling variety whereas Quench is an accepted brewing variety.

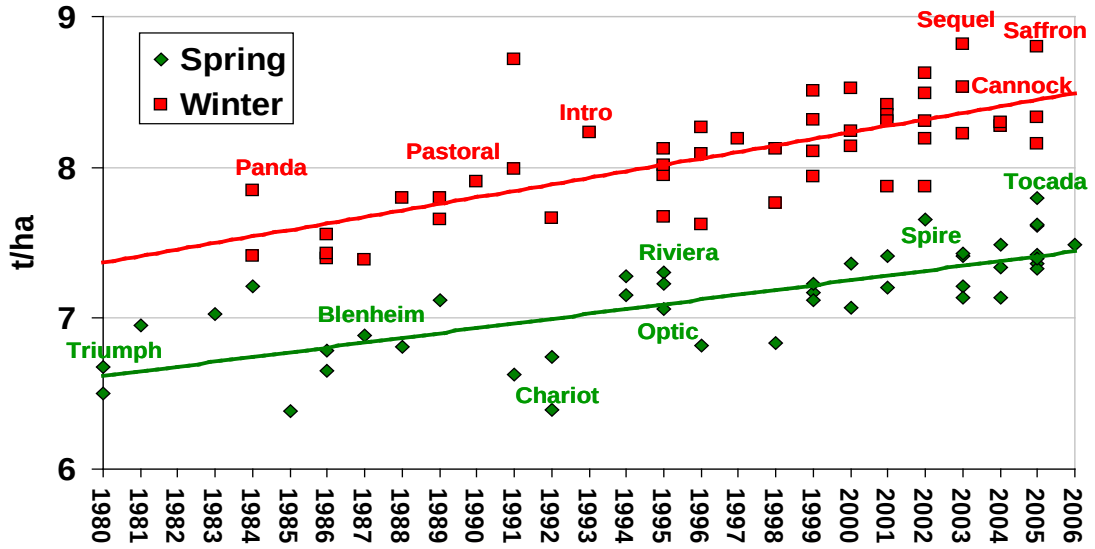



**Summary Figure 5.** Graphical genotypes of 19 spring barley varieties on the UK Recommended List for 2008. Varieties are organised in columns for chromosomes 1H to 7H. 384 BOPA1 SNP markers are organised by chromosomes into rows. Each little cell therefore represents the genotype of one variety for one marker and every one is coded with reference to Publican (right-most column for each chromosome) so that all red cells are the same as Publican and all the green cells indicate varieties carrying the alternative allele.

### 2.3.3. Genetic progress

The trials that were grown during the AGOUEB project not only provided a means of improving the prediction of means but also an unambiguous estimation of breeding progress. The varieties grown in the trials were first recommended over a range of dates between 1965 (Maris Otter) and 2006 (Appaloosa). For each crop type, we can therefore regress the mean of any variety for a given phenotype against the year in which it was first recommended and a significant relationship indicates significant genetic change over time. Summary Figure 6 shows significant improvement in yield of both the winter and spring crop over time with an average improvement of 38 kg/ha/yr for the spring and winter crop, equivalent to approximately 0.5% per annum.

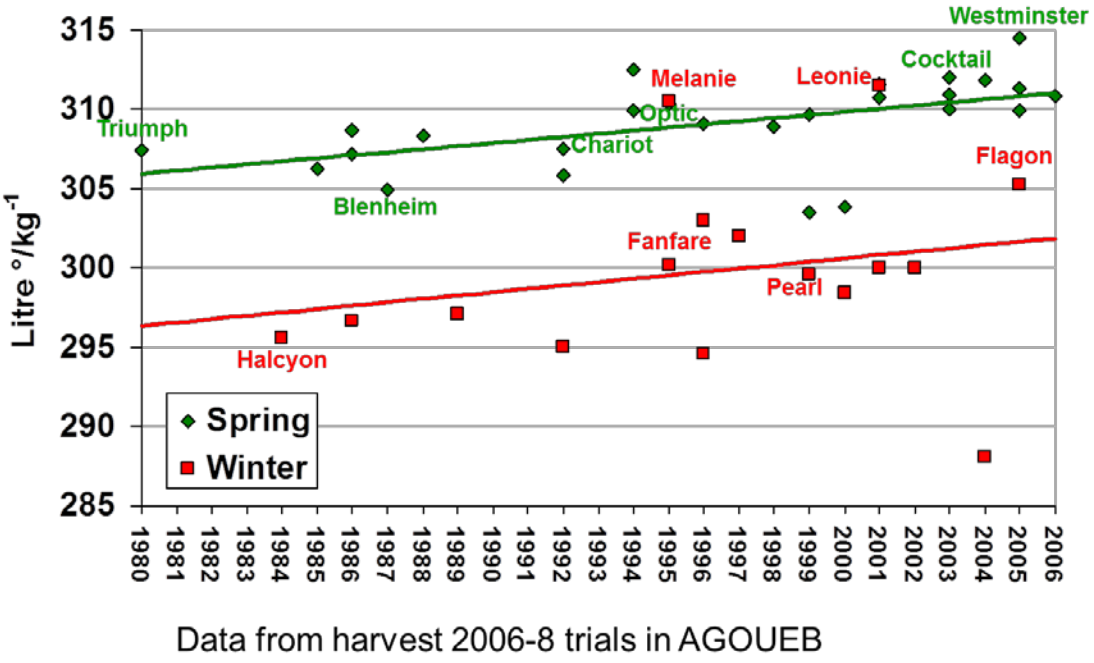
As noted in the full report, much of this increase in yield appears to be due to an increase in thousand grain weight in both crop types, which has increased by an average of 0.15g/yr over the same period. There has been no change in heading or maturity date for both crop types over the same period, indicating that improved efficiency of conversion of biomass into grain rather than an extension of the growing season was responsible for the yield increase. Despite the fact that most of the spring crop is semi-dwarf, there has been a significant decrease in plant height over the period of 0.4cm/yr but there has been no significant change in the winter crop over the same period.




**Summary Figure 6.** Overall mean fungicide treated yield derived from varieties grown together in over 20 trials from 2006-2008 plotted against their year of first recommendation. The red and green lines show the estimated linear relationship between yield for the winter and spring varieties respectively over time.

Similarly, we have also used the data derived from the analyses conducted by MAGB and SWRI to find a significant improvement in hot water extract of 0.2 litre°/kg for the spring crop over the period after the introduction of Triumph (Summary Figure 7). Whilst this does not seem large for a

character with a mean of over 300, it is highly significant for an industry consuming 1.75 million tonnes per annum in the UK. There is also an apparent increase in the hot water extract of the winter crop over the same period but this is not significant. It is biased by the apparently high extracts obtained for Melanie and Leonie, which were only assessed at one site but significantly out-performed the best control (Flagon).



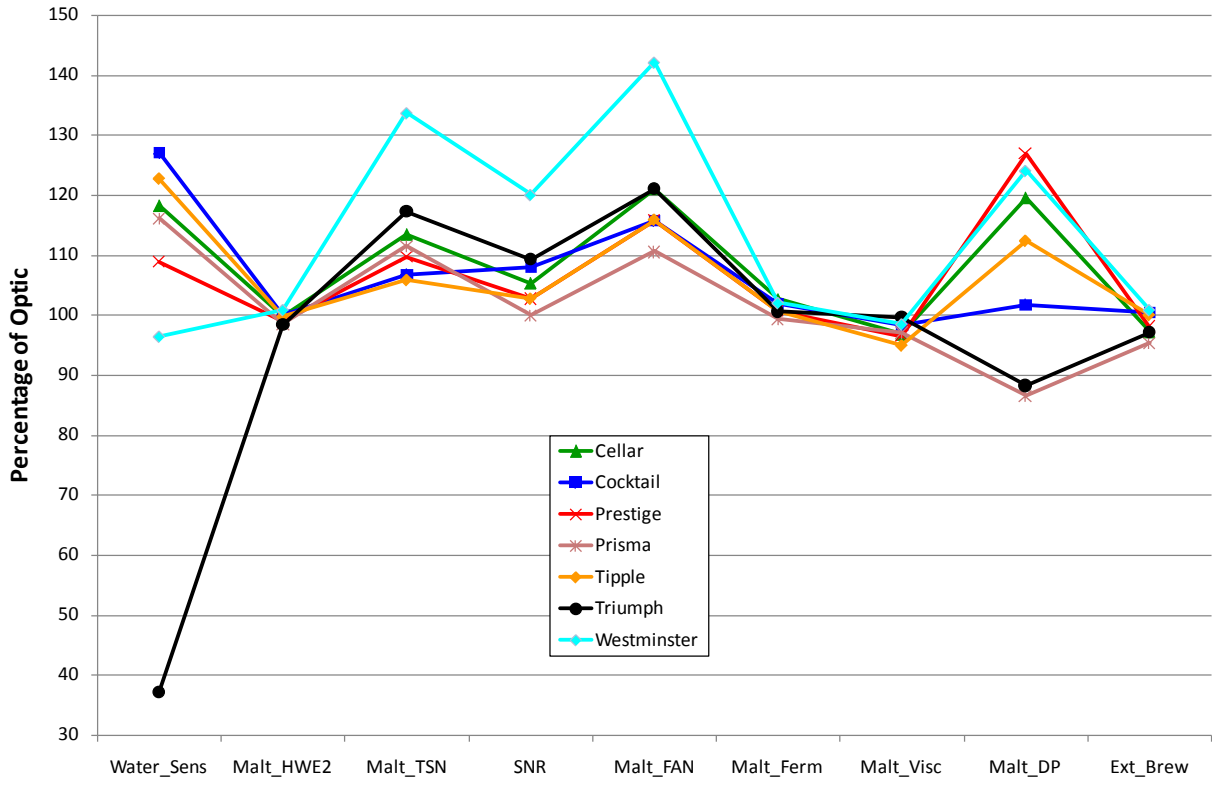

**Summary Figure 7.** Overall mean hot water extract of IBD approved spring and winter barley lines derived from analysis of micro-malted samples from trials of the varieties grown together in trials from 2006-8.

### 2.3.4. Brewing trials

We also conducted pilot brewing from 25kg lots of each of 8 spring and 4 winter barley varieties grown at two sites for each crop. Micro-malting analyses are typically carried out on small lots which may well behave differently to larger lots in commercial malting, brewing and distilling. We have analysed the data obtained from the pilot brewing to verify that genetic differences are still observable in the larger lots. All the varieties that we tested had been approved by the Institute of Brewing and Distilling so significant differences between varieties represent differences that are breeding targets to either maintain or improve upon for future varietal development.

Over 80 variates were measured on the samples from dry grain tests through to sensory attributes of the resulting brews. We detected genetic differences for 12 and 8 variates for the spring and winter barley samples respectively. Given the small number of varieties, especially for the winter crop, the finding of significant variation for some traits is highly relevant to both breeders and maltsters. It is noticeable that most of the genetic differences were found amongst the grain and processing characters and that there are few significant genetic differences for the beer and

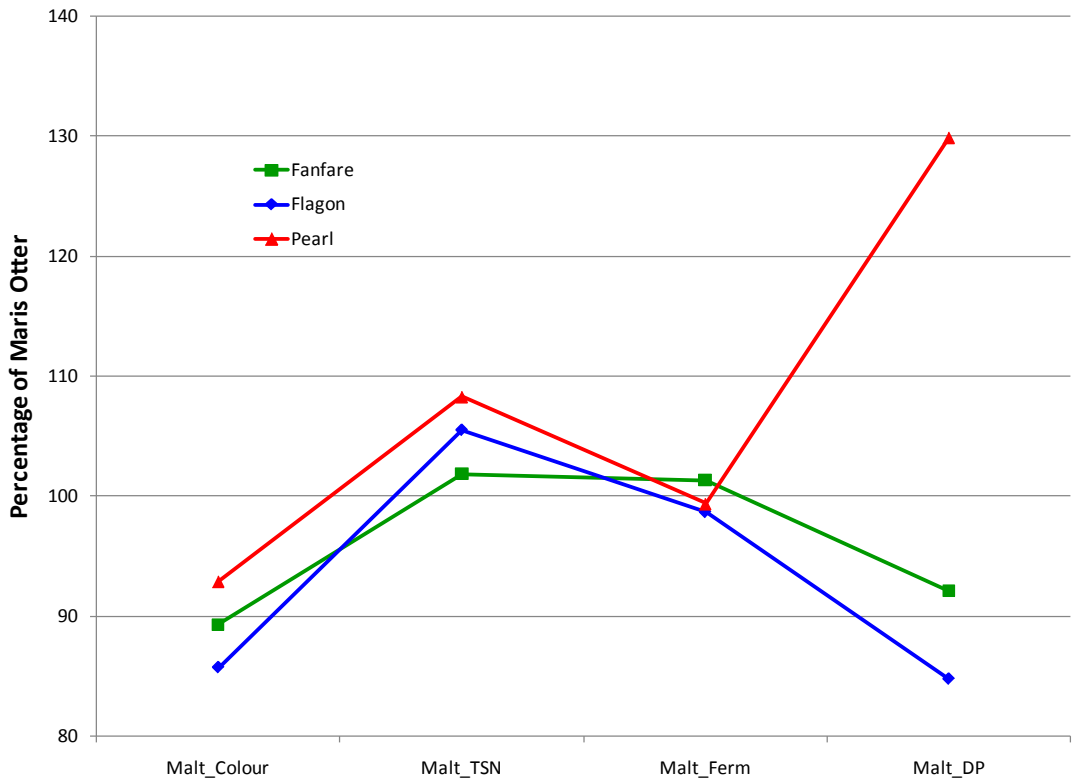
sensory characters, which illustrates the skill of the maltsters and brewers in adjusting their processes to achieve a uniform product. Amongst the springs, there was significant genetic variation for Di-methyl Sulphide (DMS) in the beer produced with the varieties Prisma and Triumph being significantly higher than the others. In general, Optic was one of the lower scoring varieties for the significant characters (Summary Figure 8), although this would partly reflect its reputation for low grain nitrogen accumulation for characters like Malt Total Soluble Nitrogen and Malt Free Amino Nitrogen. The newer varieties like Westminster and NFC Tipple tend to out-perform the older varieties like Prisma and Triumph, confirming that the genetic progress detected by micro-malting assays is reflected in an environment more akin to commercial processing.




**Summary Figure 8.** Means of 8 spring barley IBD approved varieties for nine characters relating to malt and brewing quality where there were significant genetic differences. All means are expressed as a percentage of Optic (100%).

The newer varieties have a significantly lower colour score than Maris Otter and tend to solubilise more nitrogen, whilst Pearl has a substantially higher diastatic power (Summary Figure 9). Other genetic differences, apart from three sensory attributes are, however, non-significant. Whilst not presented here, analysis of the lautering performance of Maris Otter shows that it processes remarkably well, even from a high nitrogen sample (>2%). By contrast, at least one of the two samples of the other three varieties, which were grown at the same sites, required raking for lautering to complete. Maris Otter clearly possesses attributes that have been recognised by

maltsters and brewers to the extent that some 5% of the tonnage purchased by English maltsters from harvest 2009 was Maris Otter.




**Summary Figure 9.** Means of 4 winter barley IBD approved varieties for four characters related to malt and brewing quality where there were significant genetic differences. All means expressed as a percentage of Maris Otter (100%).

## 2.4. Key conclusions

1. DNA fingerprinting has enabled the visualisation of the relationships between elite UK barley lines and their correct classification into morphological groups. Within these groups, two row malting winter barleys are clearly separated from two row feeding types.
2. We have identified strong marker trait associations for a range of DUS characters that for the first time provide an accurate view of the genomic regions sampled by such tests. The genetic resolution of the portion of the genome where the candidate genes for these characters is much finer than was possible by conventional bi-parental mapping and has permitted the identification of the major gene controlling anthocyanin pigmentation in various barley tissues.
3. Similarly, we have localised genomic regions controlling characters such as yield to a small number of candidate genes in several cases that provide real opportunities to not only utilise the associated markers in effective selection programmes but also to identify the causal genes.

4. We have also demonstrated that, despite the breeding concentration upon crosses within the elite gene pool, there remains a vast untapped reserve of unexploited genetic combinations, even when considering just recommended varieties.
5. Exploitation of new genetic combinations has sustained continued genetic progress in breeding for yield and is likely to continue to do so for the foreseeable future. The challenge is to design protocols to break down deleterious linkage blocks and re-assemble them in a targeted manner.
6. Whilst all current winter barley malting cultivars in the UK can be traced back to Maris Otter through their pedigrees, some of the processability of Maris Otter appears to have been lost in the improvement of other malting characteristics. The identification and genetic localisation of such a desirable characteristic unfortunately remains limited by the current lack of the ability to phenotype a large number of different samples, unless an appropriate surrogate test can be designed.
7. The genetic fingerprints and associated seed resources for the AGOUEB Public Set of genotypes form a biological and data resource that can and is being exploited in other projects, provided a cost-efficient means of identifying marker trait associations. For instance, projects on nitrogen use efficiency, Ramularia resistance, Fusarium resistance and second generation biofuels have all been funded and are exploiting resources generated within AGOUEB.