Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Random Forest-based genomic prediction allows selection of barley genotypes from the variety collection for inclusion evaluation in distinctness, uniformity, and stability (DUS) trials

Tally I.C. Wright
tally.wright@niab.com

Niab

Vanessa E. McMillan
Niab

Nastasiya F. Grinberg
Niab

Richard Horsnell
Niab

Malcolm Macaulay
James Hutton Institute

Luke Ramsay
James Hutton Institute

Paul D. Shaw
James Hutton Institute

James Cockram
Niab

Margaret D. Wallace
Niab

Research Article

Keywords: Variety registration, genomic prediction, barley, Random Forest, machine learning, association mapping, distinctiveness

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9584886/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Random Forest-based genomic prediction allows selection of barley genotypes from the variety collection for inclusion evaluation in distinctness, uniformity, and stability (DUS) trials

Tally I.C. Wright[1†*], Vanessa E. McMillan[1†], Nastasiya F. Grinberg[1], Richard Horsnell[1], Malcolm Macaulay[2], Luke Ramsay[2], Paul D. Shaw[2], James Cockram[1✇] and Margaret D. Wallace[1✇].

[*] Corresponding author: tally.wright@niab.com
[†] These authors contributed equally to this work and share first authorship
[✇] These authors contributed equally to this work and share last authorship
[1] Niab, Park Farm Campus, Villa Road, Histon, Cambridge, CB24 9AT, United Kingdom
[2] James Hutton Institute, Dundee, United Kingdom

Key message: Random Forest-based genomic prediction can be implemented in barley candidate variety distinctiveness assessment to enable the selection of non-distinct varieties from the variety collection for field assessment one year earlier than the current approach.

Abstract: Shortening the time taken to release a new crop variety could improve the responsiveness of the plant breeding sector. To gain Plant Breeders' Rights, and for certain marketing requirements, candidate varieties must be distinct from all known varieties in common knowledge (the variety collection) through assessment against Distinctness, Uniformity and Stability (DUS) criteria. The assessment uses field comparisons of morphological DUS characteristics between the candidate and its most similar varieties from the variety collection. Similar varieties cannot be identified immediately as candidate varieties are received without characteristic data. Here, we investigated if genomic prediction could facilitate most similar variety selection. A barley training population was assembled with 1,171 genotypes, covering ~70% of the UK barley variety collection at the time of study. Genotyping was undertaken to obtain 36,736 genetic markers and phenotypes were compiled for 28 DUS characteristics. Genomic prediction was completed using Random Forest (RF) for classification and Ridge Regression Best Linear Unbiased Prediction (rr-BLUP). RF typically showed the highest proportion of correctly classified calls and marginally higher prediction accuracies for characteristics with high heritability and fewer characteristic notes. Prediction accuracies varied more between methods for characteristics with lower heritability and more notes, where rr-BLUP performed best. Weighting mapped genetic loci in the RF prediction improved prediction accuracies. Comparisons made with historic data showed strong overlap between the most similar varieties identified via predicted or observed characteristics, indicating distinctness comparisons could begin a year earlier. Therefore, the combination of biomolecular technology and genomic prediction could enable faster distinctness assessment.

Keywords: Variety registration; genomic prediction; barley; Random Forest; machine learning; association mapping; distinctiveness.

Introduction

To help address the global challenges facing sustainable intensification of food production in a changing world, there is a need to accelerate the time taken for new plant varieties to reach farmers' fields. Conventional breeding approaches can take over 10 years to develop a new candidate variety (Voss-Fels et al., 2019). The variety listing process, required in many countries before new varieties can be marketed and sold, can add an additional 2 to 4 years to this timeline, depending on the crop and country of release (Jamali et al., 2020). Distinctness, Uniformity and Stability (DUS) testing is a key component of the variety listing process for most agricultural crop species (alongside Value for Cultivation and Use testing). In addition, DUS testing is required where a breeder has applied for variety protection via the awarding of Plant Breeders Rights (PBR). This is a form of intellectual property rights for new crop varieties that allows breeders to recoup on their investment via future royalty collection. The overall framework for plant variety protection is coordinated by the International Union for Protection of New Varieties of Plants (UPOV), an intergovernmental organisation which promotes plant variety development for the benefit of society and harmonises DUS testing procedures between member countries.

The aim of DUS testing is to use a suite of DUS 'characteristics' to establish whether a new candidate variety can be classified as phenotypically distinct from all other varieties in common knowledge (the 'variety collection'), and both uniform and stable in the expression of these characteristics (Jones et al., 2013). The variety collection can be very large for a given crop species and includes all varieties which are listed, marketed or protected by the member country, alongside varieties from other territories with similar climatic and growing conditions (UPOV, 2008).

In the United Kingdom (UK), barley DUS testing is carried out by Niab, on behalf of the Animal and Plant Health Agency (APHA). Testing is undertaken according to UPOV guidelines (UPOV, 2025) and National Protocols (APHA, 2014). Currently, there are 29 routinely recorded DUS characteristics, two special category protein electrophoresis characteristics and three approved additional characteristics included within the UK DUS test protocol for barley (Great Britain and Northern Ireland Variety Lists & UK Plant Breeders' Rights Technical Protocol for Official Examination of Distinctness, Uniformity and Stability, dated April 2025). The minimum duration of tests is normally two independent growing cycles. In the first growing cycle, a provisional description for each new candidate variety is generated based on phenotypic observations of the DUS characteristics by DUS technical experts. This data collected in year 1 is compared to the technical DUS profiles of all existing varieties in the variety collection to identify the most similar varieties to include for direct comparison in the second growing cycle using the 'similar variety selection' approach. The approach uses 'class widths', which are set and adjusted by DUS technical experts, to take into account the expected level of environmental variation and consistency of observation for individual characteristics (typically 1 note for qualitative characteristics and ≥ 3 for quantitative characteristics). The selection works on a characteristic-by-characteristic basis with varieties from the variety collection excluded from the growing trials where they are considered sufficiently different on the basis of class widths from all candidate varieties under test. When a variety from the variety collection is not considered sufficiently different then that variety will be selected for inclusion in the similar variety subset. Candidate varieties and their similar variety subsets can then be directly compared in the second growing cycle. Candidate varieties which are not yet clearly distinct on the basis of note differences in the second growing cycle can go forwards for further testing via protein electrophoresis or a third year

of side-by-side closer comparison in the field. Often a third growing cycle is required to establish distinctness for spring barley candidates due to an increasing restriction of phenotypic diversity compared to the winter barley genepool (Yang et al., 2021).

Recent advances in genomic technologies have fuelled the availability of genomic markers for use in crop breeding and research. While DUS is focused primarily on phenotypic observations, exploration of the potential utility of molecular markers in DUS processes is promoted under two main UPOV application models (UPOV, 2020). The first is like-for-like replacement of marker(s) in place of phenotypic observation of DUS characteristic(s) (UPOV, 2020). Examples in use include gene-specific markers for disease resistances in tomato (UPOV, 2024) and a male sterility characteristic in kale (UPOV, 2023). The second model involves the use of molecular distances/genetic similarities to manage the variety collection and growing trials (UPOV, 2020). For barley, the first approach alone is not feasible for all DUS characteristics due to a lack of diagnostic markers for those characteristics with low heritability and/or complex genetic architecture (Cockram et al., 2010, 2012; Yang et al., 2021). However, the characteristic-specific model could be used in conjunction with distance-based models to improve the outcomes for management of the variety collection.

Across the rest of the crop breeding and research sector, genetic markers are now routinely implemented for methods such as marker assisted selection (MAS, Hasan et al., 2021), genomic selection (Hickey et al., 2017), and quantitative trait loci (QTL) mapping in large genetic mapping populations (e.g. Wright et al., 2024). Since initial applications at the start of the 21st century (Meuwissen et al., 2001), genomic selection is now routinely implemented in commercial seed breeding programs for selection of quantitative traits (Voss-Fels et al., 2018). Genomic prediction involves the prediction of an estimated trait value for a genotype using genome-wide genetic markers, whilst genomic selection is an implementation of such predictive models in breeding programs. Once prediction models are 'trained' with an established training population (a large collection of genotypes with both genotype and phenotype data), the trait values for genotypes with only genetic data can be estimated without the need for field assessment, allowing for potential cost reduction and faster trait selection. Typical approaches to genomic prediction include parametric linear methods, such as Ridge Regression, Ridge Regression Best Linear Unbiased Prediction (rr-BLUP), and Genomic-BLUP (Bernardo and Yu, 2007; Meuwissen et al., 2001; VanRaden, 2008; Whittaker et al., 2000). These approaches have been implemented in the past using a 3k single nucleotide polymorphism (SNP) genotyping array to show varied prediction accuracies for barley DUS characteristics (Jones et al., 2012). Non-parametric machine learning tree-based approaches such as Random Forest (RF) have received recent research attention as an effective tool for prediction with genomic data (Chen and Ishwaran, 2012; Montesinos López et al., 2022). RF is robust for continuous, categorical or binary data, and under certain scenarios, has been shown to outperform linear or Bayesian-based approaches to prediction (Fradgley et al., 2023; González-Recio and Forni, 2011; Naderi et al., 2016). Furthermore, there is evidence that RF prediction accuracy or stability can be further optimised by weighting or hyperparameter tuning (Ghosh and Cabrera, 2022; Lange et al., 2025). Several previous genome-wide association studies (GWAS) have shown that some barley DUS characteristics are individually controlled by a few major-effect loci (e.g. Cockram et al., 2010; Komatsuda et al., 2007; Yang et al., 2021), weighting molecular markers linked to these regions could be a route to improved prediction accuracy (Meuwissen et al., 2024; Zhang et al., 2014). Despite the increasing availability of approaches to genomic prediction and the recent availability of genotyping arrays capable of genotyping

molecular marker numbers an order of magnitude greater than that explored by Jones et al. (2012), genomic prediction is still not implemented in the DUS testing system. Recent proof-of-concept studies have highlighted this shortfall, such as Roberts et al. (2025), who proposed a framework for how genomic prediction of quantitative characteristics could be implemented in DUS trial planning and distinctiveness assessment using perennial ryegrass and wheat as example crops.

For barley, the first growing cycle of the current UK DUS pipeline is used to assess DUS characteristics for each new candidate and most similar varieties comparisons only begin in the second growing cycle. With an appropriately trained model, the DUS characteristics of new candidates could be predicted, and comparisons with most similar varieties could begin in the first growing cycle. These earlier distinctness comparisons could potentially reduce the need for a third year of tests, improving the speed and efficiency of testing.

In this study we explore genomic prediction-based approaches focused on prediction of individual barley DUS characteristics in order to select similar varieties from the variety collection for comparison with candidate varieties in the first growing cycle. The successful implementation of a genomic prediction-based approach will rely upon the accuracy of the predictions, and this study aimed to optimise prediction methodologies for marker-based discrimination of barley varieties within the variety collection.

Methods

Barley germplasm

From the UK barley variety collection held at Niab, 537 barley varieties were selected for genotyping based on seed availability, completeness of phenotypic information, and excluding those that had been previously genotyped via the Biotechnology and Biological Research Council (BSSRC) project IMPROMALT (grants BB/K008188/1, BB/K007025/1). For each variety, a single seed was sown and grown in a temperature-controlled glasshouse at Niab (Cambridge, UK), leaf material harvested, and DNA extracted using a modified Fulton et al. (1995) protocol. DNA quality and quantity were subsequently assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific, MA, USA).

Genotypic data

SNP genotyping was carried out using the 50k Illumina Infinium iSelect barley array (Bayer et al., 2017) via subcontract to TraitGenetics GmbH (Gatersleben, Germany). The raw intensity data files (IDAT files) were extracted for 565 samples, formed from the 537 genotypes and 39 controls (negative controls and technical replicates). Early stage quality control resulted in three poor quality genotypes and eight negative controls being excluded from all downstream analyses. The remaining data were combined with the IDAT files of 749 varieties previously genotyped on the same array (Looseley et al., 2020). Following Arrieta et al. (2021), a custom SNP calling pipeline implemented in Go and R was used to score SNPs across the combined dataset (https://github.com/cardinalb/go50k).

Further quality control was completed using R (R Core Team, 2024) on the combined dataset, excluding genotypes outside the scope of the present study, technical replicates, or replicated across the two data sources. Monomorphic markers were excluded, and the SNP calls were

converted from nucleotide calls to a numerical coded format. First, molecular markers that had greater than 3% heterozygous or missing calls, or fewer than 12 homozygous calls for the minor allele were removed from the dataset. Similarly, genotypes with over 3% heterozygote marker calls, or over 1% heterozygosity and more than 0.2% missing data (the tail of the missing data distribution based on visual inspection) were removed. A Pearson's correlation test was used to compare genetic correlation for every pair of genotypes; pedigree records were inspected for each pair with a Pearson's correlation coefficient (r) > 0.95 and genotypes with unexplained relationships were excluded.

All downstream analyses were undertaken in R using the 'Crop Diversity High Performance Computing' cluster (Percival-Alwyn et al., 2024). Some R scripts were optimised using ChatGPT (OpenAI; typically, GPT-4-class models), all incorporated suggestions were reviewed and tested appropriately. Using the high-confidence and unique genotypes, marker quality control was repeated with the previously described thresholds, with the exception that any marker with missing calls was taken out in this step (the data contained relatively low amounts of missing calls). Residual heterozygosity was then marked as missing data points and then imputed as homozygote calls using the 'missForest' function implemented via the R package missForest (Stekhoven and Buehlmann, 2012). For the imputation, 150 trees were grown in each forest, the number of iterations was limited to five, and parallelisation was implemented via the 'variables' approach using 60 Central Processing Units (CPUs).

DUS phenotypic data

Historical DUS phenotypic data for all genotypes were sourced from the DUS database and paper archives maintained at Niab. The phenotypic data was in the form of DUS 'variety profiles' derived from data collected by Niab during the course of DUS examinations and variety collection management. Most variety profiles are based on two years of trial data collected during the course of the UK DUS examination. Phenotypic profiles of variety collection genotypes which were not DUS tested in the UK (varieties which have been DUS tested by other authorities) are based on one year of observations under UK conditions and have been verified, where possible, against their official variety descriptions.

The DUS data collected has changed over the period covered, as governed by changes to the UK national protocol. The number of characteristics collated reflects the 28 DUS characteristics from the UK national protocol dated 2012 and the CPVO technical protocol (CPVO-TP/19/4). Data collected from different protocol versions were converted, where necessary, to allow comparison between varieties, these conversions and several additional edits are detailed in Supplemental Text S1. The DUS characteristic number 95 (Plant: intensity of green colour) had a high proportion of missing data and was excluded from the analysis. Finally, the phenotype data were reduced to genotypes present in the genetic marker data, and genotypes with more than five missing DUS characteristics were removed.

Population structure and distance matrices

The genetic marker data were reduced to those genotypes with phenotypic observations, and markers were again filtered for minor allele count using the previously described threshold. To reduce marker redundancy, duplicated genetic markers were excluded if they had an absolute r of 1 with any other marker. The genetic marker and phenotype data were scaled prior to calculating

Euclidean distances, which were calculated using the 'dist' function in base R. Manhattan distances were calculated via the same function. An additional set of phenotypic Manhattan distances were calculated by normalising the phenotypes using min-max rescaling to a range of 0 to 1. All Manhattan distances were then standardised by dividing by the total number of genetic markers or DUS characteristics. A principal coordinate analysis (PCoA) was completed on each Euclidean distance matrix using the R package ape (Paradis and Schliep, 2019). Mantel tests with 999 permutations were used to test for Pearson's correlation between the genetic and phenotypic Manhattan matrices using the 'mantel' function from the R package vegan (Oksanen et al., 2025).

DUS characteristic heritability

As the amount of missing data varied for each DUS characteristic, for each analyses hereafter, the genetic marker matrix was reduced to only contain genotypes with phenotypes for each characteristic and genetic markers were removed that either had low minor allele counts (< 12) or were duplicated with others (an absolute $r = 1$).

Using the R package sommer (Giovanny, 2016), narrow-sense heritability (h^2) was estimated for each DUS characteristic using the 'mmes' function, where an additive relationship matrix estimated via the 'A.mat' function was treated as a random effect to estimate additive genetic variance (σ_A^2). Total phenotypic variance (σ_P^2) was estimated as $\sigma_P^2 = \sigma_A^2 + \sigma_e^2$ where σ_e^2 was the residual variance and h^2 was then estimated as $h^2 = \sigma_A^2/\sigma_P^2$.

Association mapping

Physical positions for the array markers were downloaded from the Barley 50K SNP Platform (www.ics.hutton.ac.uk/50k) and based on the Morex v1 reference genome assembly (Mascher et al., 2017). Unmapped markers were excluded from the analysis and association mapping was performed using an additive model using the R package GWASpoly (Rosyara et al., 2016), with relatedness between genotypes corrected for via a single kinship matrix estimated via the 'set.k' function. For each characteristic, a genome-wide significance level was set using the Bonferroni correction ($\alpha = 0.05$). Initially, genome-wide association study (GWAS) hits were assigned to discrete genetic loci by taking the most significant marker across a window size of 15 Mb. From these, 'major genetic loci' were then designated where $-\log_{10}(P) > 10$. Infinite *P*-values were capped at the machine minimum $-\log_{10}(P) = 307.6527$. DUS characteristic nos. 86 (Grain: type) and 88 (Grain: disposition of lodicules) were excluded from the association mapping and all downstream genomic prediction due to low phenotypic diversity. Known genetic loci and candidate genes located close to the 'major genetic loci' identified here were identified via comparison of their locations in the barley reference genome.

Genomic prediction

Twenty repetitions of 10-fold cross-validation were used for the prediction of each DUS characteristic. With the characteristics treated initially as numeric data, in each iteration the function 'createFolds' was used to stratify and partition the complete training data into ten folds using the R package caret (Kuhn and Max, 2008). For each prediction approach, the prediction accuracies and the proportion of cases correctly classified were averaged from the ten folds for each iteration. The DUS characteristics were scored with either ordinal integers or as binary (see Table 1). Two underlying approaches to genomic prediction were compared for each DUS characteristic, the first was ridge regression-BLUP used as a baseline approach with the phenotypic

data treated as numeric. The second was Random Forest (RF) classification where characteristic data were treated as distinct factors. Several methods were used for the RF approach, including hyperparameter tuning and major loci weighting with different severity.

Marker scores were recoded, the '0' scores were recalled as '-1', resulting in a dataset of '-1' and '1' representing each homozygote allele. rr-BLUP was implemented via the R package rrBLUP (Endelman, 2011). Each fold was treated as the 'test data' and the remaining 'training data' used for predicting marker effects via the 'mixed.solve' function, where the marker data was treated as a design matrix for the random marker effects. To obtain predicted characteristics for each test set, the test set marker matrix was multiplied by the estimated random marker effects, and the grand mean then added to convert the predictions to the original scale. Prediction accuracies were estimated by calculating the *r* between the test set observed and predicted phenotypes. As the predictions were no longer integers, a second prediction accuracy was recorded by rounding the predictions to the nearest integer before calculating *r*. The rounded integers were used for estimating the PCCC for the rr-BLUP prediction method.

Using the same generated folds in each iteration, the phenotype data were converted to factors and a Random Forest (RF) approach to genomic prediction for classification of DUS characteristics was implemented using the randomForestSRC R package (Ishwaran and Kogalur, 2025). Forests were initially grown using the default arguments of the 'rfsrc' function and the test data predicted characteristic notes were obtained. Following methods outlined by Montesinos López et al. (2022), a hyperparameter inner-fold tuning approach was implemented to explore the influence of changing the number of trees grown per forest (*ntree*), the minimum node size (*nodesize*) and the number of variables considered at each split (*mtry*). The values of *ntree* considered were 250, 500 (default), and 750. For *nodesize*, values of 1 (default), 3, and 6 were tested. For *mtry* the default was equal to the square root of the number of predictor variables (genetic markers). Where phenotype data were available for all varieties, there were 18,545 unique genetic markers which corresponded to a default *mtry* value of 137. The values of *mtry* tested were 50, 150, 250, and 350. A full factorial comparison was used, resulting in 36 combinations of hyperparameters. The inner-fold tuning approach involved drawing five inner-folds from the training population of each outer-fold. For each inner-fold, inner test values were predicted by forests trained using each of the 36 combinations of hyperparameters. The combination that produced the highest PCCC was then used as the hyperparameters for predicting that outer-fold test data. This process was repeated for each outer-fold within each iteration.

The final prediction model used the default hyperparameters without inner-fold tuning but increased the weight of those prediction variables (genetic markers) identified as major genetic loci in the association mapping. This was achieved by supplying a vector of marker weights to the arguments *xvar.wt* (controls the probability of selecting a genetic marker for possible splitting of a node) and *split.wt* (controls the split statistic of the selected genetic markers) of the 'rfsrc' function (Ishwaran and Kogalur 2025). For characteristics with 'major genetic loci' identified, two weighting levels were tested. All markers were initially assigned a weight of '1' (i.e. no preferential treatment of any marker) and then in the first approach peak genetic markers associated with major genetic loci were weighted as '2' (i.e. the weighting controlling the probability of selecting candidates for splitting and the split statistic is increased twofold), in the second approach the same markers were weighted as '10' (i.e. the weighting controlling the probability of selecting candidates for splitting and the split statistic is increased tenfold).

To test the effect of prediction method on prediction accuracy or PCCC, iteration was treated as a blocking factor, and linear mixed models were fitted for each trait as:

$$y_{ij} = \mu + m_j + r_i + e_{ij}$$

where y_{ij} was the prediction accuracy or PCCC for the prediction method j with the i iteration, μ was the grand mean, m was prediction method treated as a fixed effect, r was iteration treated as a random effect and e was the residual error. The models were fitted and analysis of variance conducted for the fixed effects using the R packages lme4 and lmerTest (Bates et al., 2015; Kuznetsova et al., 2017). Estimated marginal means and standard errors for the prediction methods were calculated using the R package emmeans (Lenth, 2024) and compact letter display for pair-wise comparisons with P-values adjusted via the Tukey method were calculated using the 'cld' function from the R package multcomp (Hothorn et al., 2008).

Similar variety selection

A custom script was created in R that replicated the 'similar variety selection' approach used in current UK barley DUS system. The most recent version of class widths for the characteristics was used (Supplementary Table 1). For seasonal type (UK national characteristic no. 90), the class width was reduced from '2' to '1' to reflect changes to the characteristic scoring (Supplementary Text 1). Registered barley varieties present in the dataset released between 2010 and 2019 were treated as candidate varieties. For each year of entry, these registered varieties were treated as the candidate set and their characteristics were predicted using the rest of the genotypes in the panel as the reference population. The prediction was completed through a classification approach using the randomForestSRC R package with default settings (Ishwaran and Kogalur, 2025). For each candidate, a comparison was then made between the subsets of variety collection varieties identified as non-distinct and thus classified as required for field evaluation when using either observed or predicted characteristic data for the candidates. Subset size using observed or predicted data was visualised with the R package ggplot2 (Wickham, 2016). Excluding candidates with one or fewer variety in the non-distinct subset identified with observed data, the percentage of variety collection genotypes identified as non-distinct using observed phenotypic data also found with predicted phenotypes was estimated.

Results

The barley prediction panel

SNP calling was completed with a combined set of 565 barley samples genotyped in this study on the UK barley variety collection held at Niab and 749 previously genotyped barley samples (Looseley et al., 2020). After excluding controls, replicates and low-quality samples, 1,171 unique genotypes were used for the study, which included 459 and 712 from the recent and previous genotyping efforts, respectively. Of the 43,799 markers outputted initially from the SNP calling pipeline, 36 markers were removed due to formatting uncertainty, 1,609 excluded as monomorphic, 1,366 markers removed with missing data calls, 2,932 markers removed with heterozygosity >3%, and 1,120 excluded with <12 calls of the homozygous minor allele. The resulting dataset contained 36,736 quality controlled genetic markers, of which 16,540 contained

residual heterozygosity (<3%). These heterozygous calls were marked as NA and imputed as homozygous alleles.

Of the 1,171 genotypes in the combined dataset, 1,129 had DUS characteristic data. This number was further reduced by removing 42 genotypes with >5 missing DUS characteristics and a single anomaly, resulting in a data set of 1,086 genotypes with genetic and phenotypic data, hereafter referred to as the 'barley prediction panel'. There were 28 DUS characteristics scored and curated for the barley prediction panel, all scored by discrete values ranging from 1 to 9 (Table 1). Eight of these characteristics had binary notes (either '1' or '2'), including the important barley characteristics Ear: number of rows (characteristic no. 39) and seasonal type (no. 90). For characteristic no.39 (Ear: number of rows), there were 1,000 two-row (note '1') and 86 six-row (note '2') barley genotypes present in the data. There was a more even split for seasonal type, with 421 winter (note '1') and 593 spring (note '2') types in the barley prediction panel. Limited phenotypic diversity was observed for the DUS characteristics Grain: type (no. 86) and Grain: disposition of lodicules (no. 88), and therefore these characteristics were excluded from the genetic mapping and genomic prediction. Of the non-binary characteristics, three DUS characteristics were scored with three characteristic notes and the remaining 17 were scored with scales ranging from five to nine characteristic notes. Missing data varied trait-by-trait, ranging from NA = 0 (seven characteristics) to NA = 571 (no. 9; Flag leaf: attitude) (Table 1).

Table 1 - Note frequencies and heritability (h^2) for the 28 DUS characteristics analysed in the barley prediction panel (1,086 genotypes with genetic marker and phenotypic data). The standard error (s.e.) for each heritability estimate (h^2), number of genotypes with observed data (n) and the number of missing data points (NA) for each characteristic is shown.

Characteristic no.	Name of characteristic	h^2	s.e.	1	2	3	4	5	6	7	8	9	n	NA
1	Plant: growth habit	0.32	0.04	2	7	72	186	370	236	155	41	9	1078	8
2	Lower leaves: hairiness of leaf sheaths	0.97	0.01	675	410	-	-	-	-	-	-	-	1085	1
8	Time of ear emergence	0.31	0.04	19	50	107	221	321	244	88	26	8	1084	2
9	Flag leaf: attitude	0.40	0.06	18	37	111	165	123	36	12	4	9	515	571
10	Flag leaf: intensity of anthocyanin colouration of auricles	0.72	0.03	145	5	45	50	103	185	337	86	36	992	94
14	Ear: glaucosity	0.42	0.04	33	20	66	141	365	282	158	18	3	1086	0
15	Flag leaf: glaucosity of sheath	0.14	0.03	0	0	7	16	143	265	461	171	23	1086	0
17	Ear: attitude	0.30	0.04	18	246	279	215	178	97	44	7	0	1084	2
19	Awns: intensity of anthocyanin colouration of tips	0.66	0.04	142	13	65	103	270	252	128	19	1	993	93
20	Sterile spikelet: attitude	0.64	0.05	49	121	494	-	-	-	-	-	-	664	422
22	Ear: development of sterile spikelets	0.82	0.02	365	625	-	-	-	-	-	-	-	990	96
25	Median spikelet: length of glume and its awn relative to grain	0.24	0.04	51	896	131	-	-	-	-	-	-	1078	8
29	Ear: length	0.17	0.03	1	7	52	164	514	265	66	6	0	1075	11
30	Awn: length	0.27	0.04	0	0	24	26	216	146	592	39	1	1044	42
33	Plant: length	0.21	0.04	29	67	182	318	287	134	49	14	4	1084	2
39	Ear: number of rows	1.00	0.00	1000	86	-	-	-	-	-	-	-	1086	0
40	Ear: density	0.29	0.04	10	34	226	273	396	98	46	3	0	1086	0
41	Ear: shape	0.12	0.03	-	-	494	68	480	0	1	-	-	1043	43
44	Rachis: length of first segment	0.31	0.04	0	0	63	204	574	193	51	1	0	1086	0
46	Rachis: curvature of first segment	0.27	0.04	15	54	220	255	321	141	71	4	3	1084	2
62	Kernel: colour of aleurone layer	0.76	0.03	814	91	92	-	-	-	-	-	-	997	89
65	Grain: rachilla hair type	0.93	0.01	310	776	-	-	-	-	-	-	-	1086	0
73	Grain: spiculation of inner lateral nerves	0.75	0.03	626	57	145	25	61	16	70	37	49	1086	0
75	Grain: anthocyanin colouration of lemma nerves	0.62	0.04	268	52	121	169	239	133	72	24	1	1079	7
76	Grain: hairiness of ventral furrow	0.99	0.01	934	151	-	-	-	-	-	-	-	1085	1
86	Grain: type	0.19	0.03	3	1082	-	-	-	-	-	-	-	1085	1
88	Grain: disposition of lodicules	0.99	0.01	10	1066	-	-	-	-	-	-	-	1076	10
90	Seasonal type	1.00	0.00	421	593	-	-	-	-	-	-	-	1014	72

Heritability (h^2) varied greatly on a characteristic-by-characteristic basis (Table 1). The highest heritability values (h^2 = 1) were observed for Ear: number of rows (no. 39) and Seasonal type (no. 90), while the lowest was found for Ear: shape (no. 41; h^2 = 0.12). Following definitions previously used by Yang et al. (2021), barley DUS characteristics were grouped as either having low (h^2 < 0.5) heritability (15 characteristics) or high (h^2 > 0.5) heritability (13 characteristics). Of the latter group, nine had three or less characteristic notes, while the remaining four characteristics were scored on a 1-9 scale. In the low heritability group, only two of the 15 characteristics were scored using ≤3 notes.

Principal coordinate analysis (PCoA) of Euclidean distances calculated using either the scaled genetic marker or phenotypic data showed clear population structure in the barley prediction panel (Figure 1a-b). The first principal coordinate (PCoA_1) separated the genotypes by seasonal type (no. 90) and explained 18% and 19% of the variation when using genetic or phenotypic-based Euclidean distances, respectively. The PCoA for both approaches showed separation of the genotypes by Ear: number of rows (no. 39) on the second principal coordinate (PCoA_2), where the higher values were associated with the six-row genotypes. There was less variation explained by PCoA_2 for the genetic-based Euclidean distances (4%), than phenotypic-based estimates (11%). However, the clustering of the variation associated with characteristic nos. 90 and 39 was visually clearer when using marker-based Euclidean distances (Figure 1a). The Manhattan distances estimated from the genetic markers and phenotype characteristics were significantly correlated (P = 0.001, Figure 1c-d). The correlation increased in strength when using min-max scaling to normalise the phenotypic characteristics scored with a different number of notes prior to the calculation of Manhattan distances; with normalised phenotypes there was a Pearson's correlation coefficient (r) of 0.76, compared to 0.61 without normalisation.

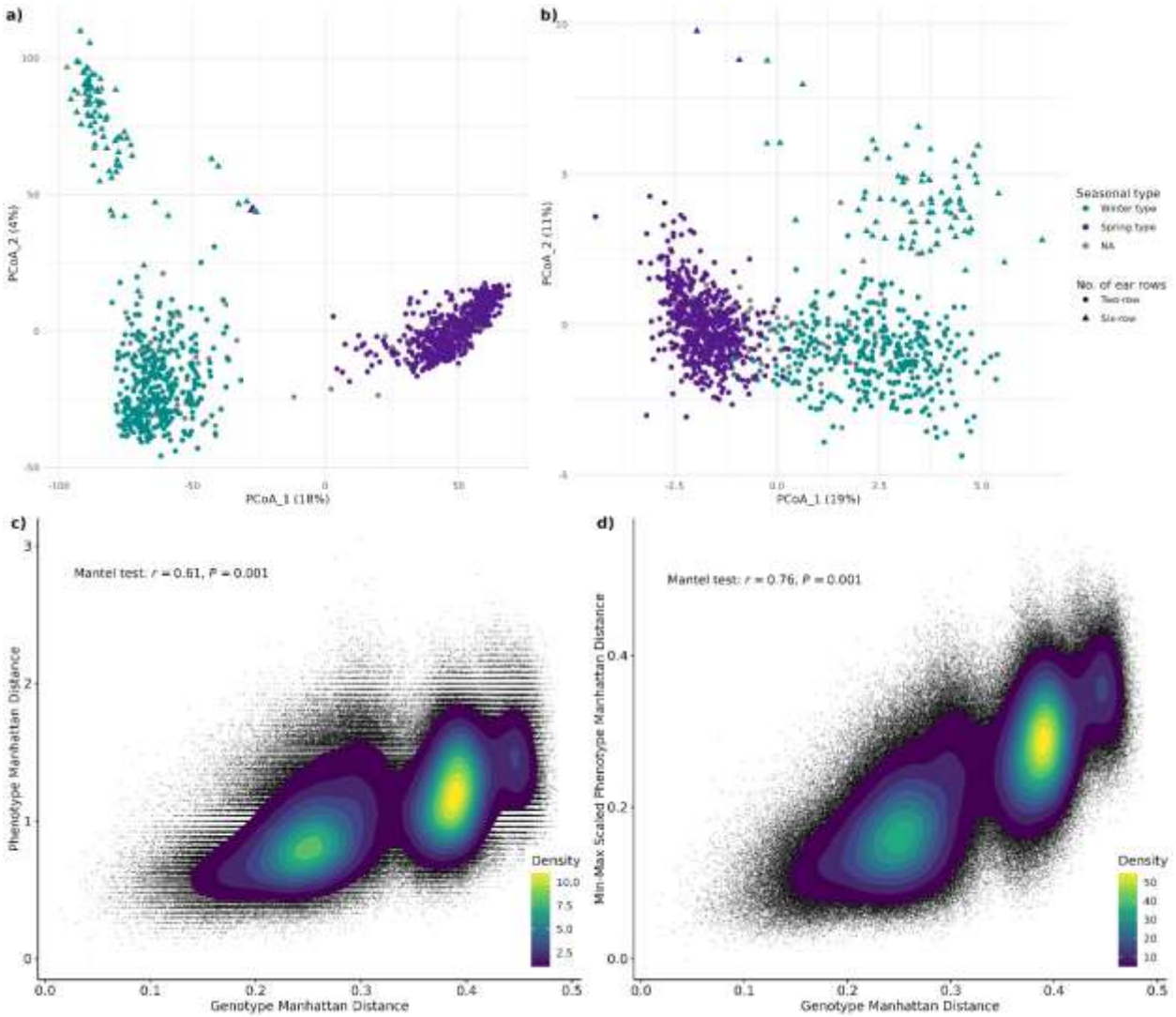


Figure 1 - Trends and correlation between the phenotypic and genotypic spaces of the barley prediction panel (n = 1,086). Population structure is illustrated using a principal coordinate analysis (PCoA) calculated using Eucladian distances estimated from either the scaled genetic marker data (a) or the scaled phenotypic DUS characteristic data (b). For each PCoA, the first two principal coordinates are shown (PCoA_1 and PCoA_2) and the percentage variation explained by each of these coordinates is shown in the axis labels. The genotypes are coloured by Seasonal type and shaped by Ear: number of rows. Correlations between the Manhattan distances calculated using genetic marker data and the phenotypic DUS characteristics before (c) and after (d) normalisation using min-max scaling. Mantel test statistics are overlaid on each plot and point density smoothing is represented by a 2D kernel density estimation for each comparison.

Association mapping

The phenotypic distributions of the 26 DUS characteristics used for GWAS are shown in Supplementary Figure 1, and the resulting Manhattan and quantile-quantile (QQ) plots are shown in Supplementary Figures 2-4. Association mapping identified 23 ‘major genetic loci’ (QTL with $-\log_{10}(P) > 10$) across 14 of the DUS characteristics (Table 2), with the $-\log_{10}(P)$ scores for all these loci above the Bonferroni genome-wide $\alpha = 0.05$ significance thresholds (Supplementary Table 2). Association mapping identified at least one QTL for each of the 12 DUS characteristics

with $h^2 > 0.5$ (Table 1). The highest number of QTL were identified for Ear: number of rows (no. 39) where six 'major genetic loci' were mapped across chromosomes 1H, 2H, 4H, 5H and 7H, and for Seasonal type (no. 90) where four 'major genetic loci' were identified on chromosomes 1H, 4H and 5H. For traits with heritability $h^2 < 0.5$, 'major genetic loci' were only found for Ear: glaucosity (no. 14) and Median spikelet: length of glume and its awn relative to grain (no. 25) (Table 2). Genetic markers associated with the 'major genetic loci' were weighted to investigate possible optimisation of the Random Forest (RF) prediction approach.

Table 2 - Summary of the 'major genetic loci' ($-\log_{10}(P) > 10$) identified via association mapping in the barley prediction panel (n = 1,086). Displayed for each major locus is the name (Marker), assigned chromosome and position of each peak marker on the barley reference genome assembly. Where relevant, candidate genetic locus, candidate gene (or SNP name in the instance of previously identified GWAS hits) and associated references are shown. * DUS characteristics with low heritability ($h^2 < 0.5$). Extended results for this table are shown in Supplementary Table 2.

Characteristic no.	Marker	Chromosome	Position (bp)	$-\log_{10}(P)$	Candidate locus	Candidate Gene/GWAS SNP	References
2	JHI.Hv50k.2016.269262	4H	631,633,338	180.92	*HSH1*	GWAS hit (SNP_708353)	Cockram et al., 2010; Cockram et al., 2012; Kim et al., 2023
10	JHI.Hv50k.2016.110602	2H	676,851,667	78.89	*ANT2*	*bHLH*	Cockram et al., 2010
14*	SCRI_RS_204276	1H	248,411	12.11	*GLS1*	GWAS hit (12_30969)	De la Fuente Canto et al., 2018
19	JHI.Hv50k.2016.110602	2H	676,851,667	94.73	*ANT2*	*bHLH*	Cockram et al., 2010
20	JHI.Hv50k.2016.28062	1H	394,166,483	16.92	*VRS3* (*Int-A*)	HvJMJ706	Bull et al., 2017
20	JHI.Hv50k.2016.108012	2H	654,157,564	41.94	*Vrs1*	*HvHOX1*	Komatsuda et al., 2007
22	JHI.Hv50k.2016.107661	2H	651,539,496	203.18	*Vrs1*	*HvHOX1*	Komatsuda et al., 2007
25*	JHI.Hv50k.2016.460699	7H	42,951,647	30.63	N/A	GWAS hit (SNP_1172437)	Kim et al., 2023
39	JHI.Hv50k.2016.16862	1H	28,824,035	10.22			
39	JHI.Hv50k.2016.44059	1H	518,171,392	14.72	*EAM8*	*HvELF3*	Faure et al., 2012
39	JHI.Hv50k.2016.107123	2H	649,821,590	307.65	*Vrs1*	*HvHOX1*	Komatsuda et al., 2007
39	JHI.Hv50k.2016.231492	4H	19,763,503	307.65	*Int-C* (*Vrs5*)	*HvTB1*	Ramsay et al., 2011
39	JHI.Hv50k.2016.288224	5H	34,677,519	10.27			
39	BOPA2_12_20793	7H	39,718,323	21.64	*VRN-H3*	*FT1*	Yan et al., 2006
62	JHI.Hv50k.2016.250677	4H	524,864,313	134.95	*BLX*	(1) *HvMyc2*; (2) *HvMpc2*	Strygina et al., 2017
65	JHI.Hv50k.2016.316002	5H	546,260,097	177.92	*SRH1*	*HvSRH1*	Jayakodi et al., 2024
73	JHI.Hv50k.2016.104839	2H	639,289,447	71.96	*GTH1*	GWAS hit (Hv11_10818_CA)	Cockram et al., 2010; Cockram et al., 2012
75	SCRI_RS_160618	2H	676,848,238	64.60	*ANT2*	*bHLH*	Cockram et al., 2010
76	JHI.Hv50k.2016.367872	6H	330,411	146.16	N/A	GWAS hit (HvOs02g01490_G607A)	Cockram et al., 2010; Cockram et al., 2012
90	JHI.Hv50k.2016.41782	1H	511,924,294	133.93	*PPD-H2*	*HvFT3*	Faure et al., 2007
90	JHI.Hv50k.2016.275319	4H	643,513,006	214.94	*VRN-H2*	*ZCCTabc*	Yan et al., 2004; Dubcovsky et al., 2005
90	JHI.Hv50k.2016.323294	5H	571,615,210	14.53	*VRN-H1*	*HvBM5A*	Distelfeld and Dubcovsky, 2009
90	JHI.Hv50k.2016.336854	5H	601,232,579	57.84			

Genomic prediction

Regardless of the prediction approach, there were strong observed trends between characteristic heritability and either prediction accuracy (r) or the proportion of correctly classified calls (PCCC) (Figure 2). DUS characteristics with higher heritability also had higher r and PCCC. The highest r and PCCC values were observed for characteristics no. 90 (Seasonal type) and no. 39 (Ear: number of rows), where all values across all methods tested >0.98 (Table 3). For 11 of the 12 DUS characteristics with high heritability ($h^2 > 0.5$), one or more of the RF approaches resulted in higher r values than ridge regression-BLUP (rr-BLUP) or with the rr-BLUP prediction rounded to the nearest integer (Table 2). Although there were significant effects of prediction method on r for all the characteristics investigated (Table 2), the increase in r values for traits with high heritability offered by the RF methods were relatively small. For example, out of the 12 DUS characteristics with $h^2 > 0.5$, the largest increase between rr-BLUP and RF with default parameters was 0.08 for Grain: rachilla hair type (no. 65). For the characteristics with high heritability ($h^2 > 0.5$), optimisation of the RF approach through tuning or weighting typically resulted in small improvements to prediction accuracy. For instance, across the 12 DUS characteristics with $h^2 > 0.5$, RF with the highest degree of 'major genetic loci' weighting (rf_weight10) had a mean r of 0.90, compared to rr-BLUP and RF with default parameters that returned mean r values of 0.86 and 0.88, respectively.

There was greater variation for r between prediction methods for characteristics with $h^2 < 0.5$ (Figure 2a), where rr-BLUP typically outperformed the RF classification approach. Here the variation was larger, for the 14 characteristics with $h^2 < 0.5$ the mean r for rr-BLUP was 0.47 compared to 0.36 for RF with default settings (Table 2). For the RF method with default settings, the lowest r was found for characteristic no. 44 (Rachis: length of first segment; $r = 0.24$). Whereas, for rr-BLUP the prediction was more accurate for the same characteristic ($r = 0.48$). For each of the two of the 'low heritability' DUS characteristics found to be controlled by 'major loci' (no .14 - Ear: glaucosity; and no. 25 - Median spikelet: length of glume and its awn relative to grain) there was an increase in r observed using 'major loci' weighting compared to the default RF approach (Table 2).

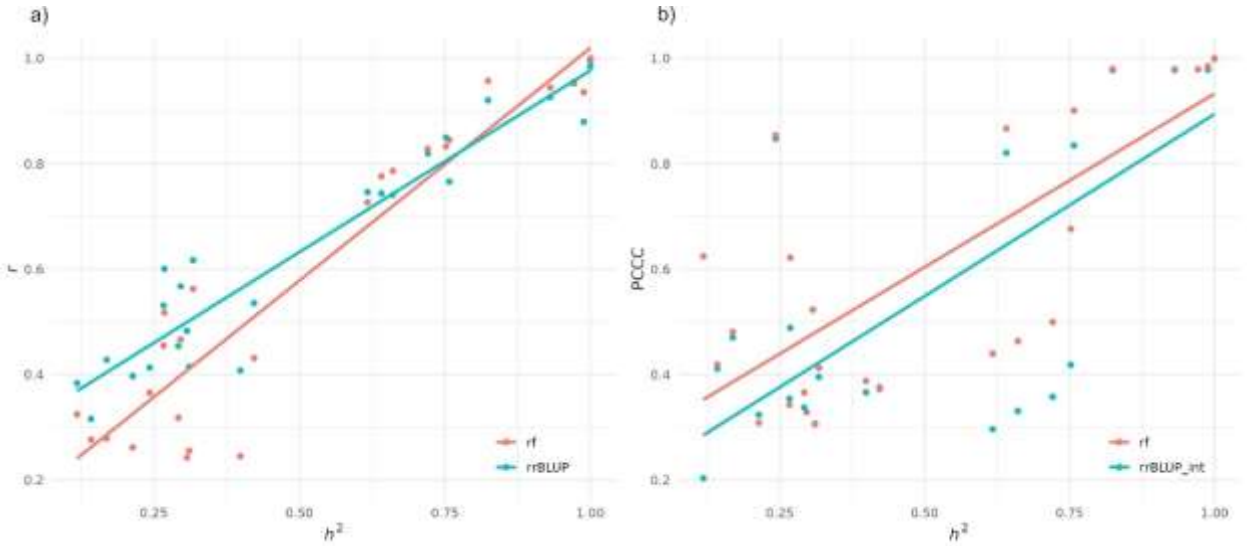


Figure 2 - Comparisons between trait heritability (h^2) and (a) prediction accuracy (r) and (b) proportion of correctly classified calls (PCCC). For either r or PCCC, each point represents an average of 20 interactions for one of the DUS characteristics predicted using either Random Forest with default parameters (rf) or Ridge Regression-BLUP with prediction rounded to the nearest integer (rrBLUP_int). Lines of best fit are shown.

Table 3 - Average prediction accuracy (r) and proportion of correctly classified (PCCC) for 26 DUS characteristics analysed in the barley prediction panel (n = 1,086). The r and PCCC values were averaged from 20 repetitions of 10-fold cross-validation and the genomic prediction methods analysed: Ridge Regression-BLUP (rrBLUP), rrBLUP predictions rounded to the nearest integer (rrBLUP_int), Random Forest (RF) with default parameters (rf), RF with hyperparameter tuning (rf_tune), and RF with different levels of genetic marker weighting (rf_weight2 and rf_weight10). The weighting approach was only used for DUS characteristics found via association mapping to be controlled by 'major loci' ($-\log_{10}(P) > 10$). Mixed linear models were fitted for each characteristic and used to test the effect of prediction method on r and PCCC for each trait, P-value signicance thresholds (Sig. level) were extracted from ANOVA tables for these comparisons (*** = $P < 0.001$; *n.s.* = $P > 0.05$). Full DUS characteristic descriptions are shown in Table 1.

Characteristic no.	rrBLUP	rrBLUP_int	rf	rf_tune	rf_weight2	rf_weight10	Sig. level	rrBLUP_int	rf	rf_tune	rf_weight2	rf_weight10	Sig. level
				r						PCCC			
1	0.617	0.580	0.562	0.556			***	0.395	0.412	0.413			***
2	0.953	0.956	0.956	0.973	0.959	0.980	***	0.979	0.979	0.987	0.980	0.991	***
8	0.415	0.384	0.256	0.247			***	0.307	0.306	0.305			*n.s.*
9	0.408	0.386	0.245	0.234			***	0.366	0.388	0.391			***
10	0.820	0.809	0.828	0.844	0.831	0.854	***	0.358	0.500	0.502	0.499	0.502	***
14	0.536	0.502	0.431	0.436	0.435	0.445	***	0.375	0.372	0.378	0.375	0.375	*n.s.*
15	0.316	0.239	0.276	0.272			***	0.411	0.419	0.417			***
17	0.568	0.544	0.467	0.467			***	0.329	0.329	0.327			*n.s.*
19	0.741	0.727	0.786	0.796	0.790	0.808	***	0.331	0.463	0.464	0.462	0.463	***
20	0.744	0.705	0.776	0.812	0.798	0.831	***	0.821	0.867	0.873	0.872	0.881	***
22	0.920	0.952	0.957	0.957	0.960	0.969	***	0.978	0.980	0.980	0.981	0.985	***
25	0.413	0.299	0.365	0.386	0.373	0.410	***	0.848	0.854	0.858	0.855	0.861	***
29	0.428	0.329	0.279	0.276			***	0.471	0.481	0.481			***
30	0.601	0.553	0.517	0.504			***	0.488	0.622	0.620			***
33	0.397	0.329	0.262	0.261			***	0.324	0.309	0.311			***
39	0.984	0.992	1.000	0.993	1.000	1.000	***	0.999	1.000	0.999	1.000	1.000	***
40	0.454	0.410	0.318	0.321			***	0.337	0.366	0.370			***
41	0.384	0.294	0.325	0.337			***	0.203	0.624	0.631			***
44	0.483	0.408	0.242	0.239			***	0.523	0.523	0.523			*n.s.*
46	0.531	0.498	0.455	0.450			***	0.354	0.342	0.345			***
62	0.766	0.738	0.845	0.859	0.847	0.856	***	0.835	0.901	0.904	0.902	0.904	***
65	0.927	0.949	0.944	0.944	0.944	0.958	***	0.979	0.978	0.977	0.978	0.983	***
73	0.850	0.842	0.833	0.846	0.837	0.848	***	0.419	0.676	0.680	0.678	0.681	***
75	0.746	0.731	0.727	0.771	0.740	0.767	***	0.296	0.440	0.449	0.444	0.447	***
76	0.879	0.909	0.936	0.928	0.934	0.938	***	0.979	0.985	0.983	0.985	0.986	***
90	0.995	0.999	0.999	0.999	0.999	0.999	***	1.000	1.000	1.000	1.000	1.000	*n.s.*

The RF approach was generally more effective for PCCC across all characteristics compared to rr-BLUP with integer rounding (Figure 2b). However, unlike the case for r, five characteristics were identified for which prediction method had no significant effect on PCCC (Table 2). For the 12 characteristics with high heritability ($h^2 > 0.5$), RF with the higher degree of weighting had the highest mean PCCC of 0.82. For these same characteristics, rr-BLUP with integer rounding had a mean PCCC of 0.75. Unlike for r, for characteristics with low heritability ($h^2 < 0.5$) the rr-BLUP method performed worse in terms of PCCC with a mean of 0.41, compared to RF with default settings which had a mean of 0.45. Notably, there was a 0.42 and 0.26 drop in PCCC using rr-BLUP with integer rounding compared to the RF default approach for Ear: shape (no. 41) and Grain: spiculation of inner lateral nerves (no. 73), respectively.

Genetic marker weighting in the RF approach offered higher increases in r and PCCC compared to hyperparameter tuning. Across all traits, tuning of the RF method resulted in a mean r of 0.60 and PCCC of 0.62, which is the same average observed for RF with default settings. Out of the two weighting approaches for RF, the more severe weighting approach (rf_weight10) resulted in the largest mean increase (r = 0.02) compared to using the default RF settings. For 11 out of the 14 characteristics that used weighting, the increase observed between the default and the most severe weighting method for r was statistically significant (Supplementary Figures 5 and 6). Overall, the genomic prediction approach that provided either the highest r or PCCC varied on a characteristic-by-characteristic basis. For each characteristic, pairwise comparisons are shown in Supplementary Figures 5-8 for both r and PCCC.

Implementing predictions

The predicted and observed DUS characteristic scores were compared, for subsets of registered varieties over a 10-year period, to the phenotypic profiles of all varieties in the dataset to identify the most similar varieties that would be required for direct distinctness assessment in the field, using the current class width approach for the selection of similar varieties. The comparisons between the genotypes selected from the variety collection using observed or predicted DUS characteristics as the basis of the selection are shown in Figure 3. The size of subsets required for field testing using the predicted characteristics typically increased with the number of varieties shown to be not yet clearly distinct using the observed data (Figure 3a). However, the average number of varieties in the prediction-based subsets (n = 49) was noticeably higher than the average number in the observed characteristic-based subsets (n = 29), suggesting an average 69% increase in the number varieties required for field testing using the prediction-based approach.

Although the prediction-based subsets were larger, there was good representation of the varieties identified as not yet clearly distinct and required for field assessment using observed characteristic data (Figure 3b). Excluding candidate varieties with one or less variety in the identified observed subset, on average 72% of varieties found to be not yet clearly distinct with observed characteristic data were also identified using predicted characteristics. There was a left-skewed distribution visible in the percentages and six varieties had no observed overlap (Figure 3b).

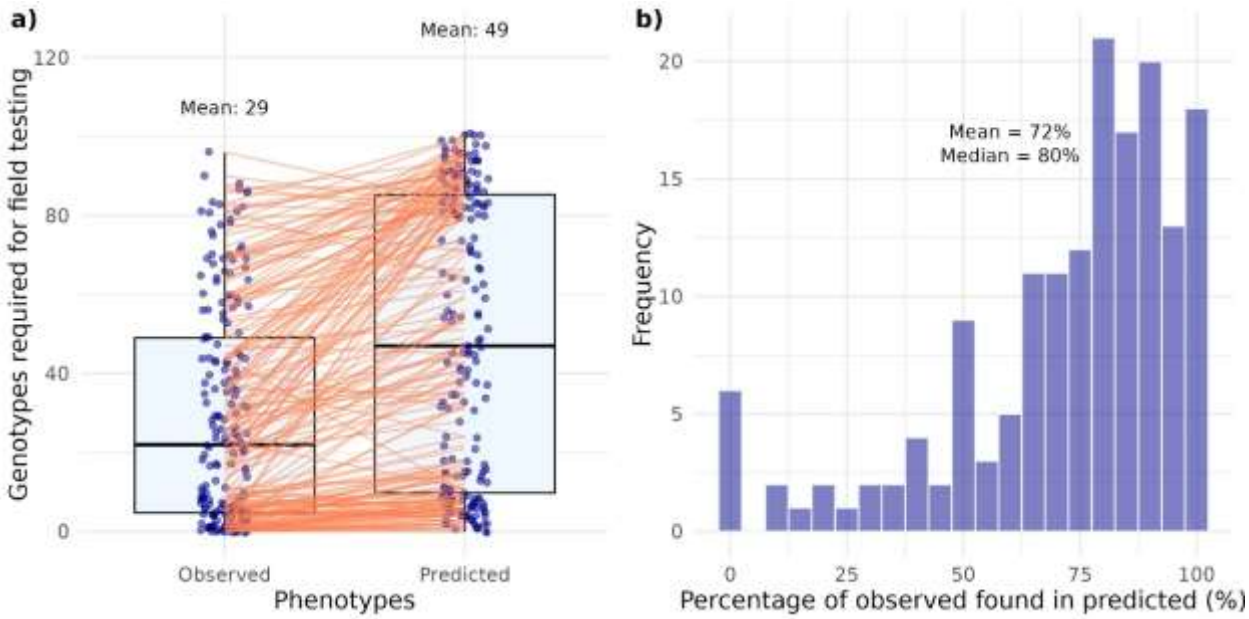


Figure 3 – Comparison of the number of genotypes from the 'variety collection' subsets shown to be not yet clearly distinct from registered varieties (treated as candidate varieties) across a ten-year period, undertaken using either observed or predicted DUS characteristics. In the current barley DUS pipeline, these subsets of the 'variety collection' would be grown in the field alongside the candidate varieties for direct distinctness comparisons. a) The number of 'variety collection' genotypes found as not yet clearly distinct and therefore required for field testing using either observed or predicted characteristics, with orange connecting lines indicating the subsets associated with the same candidate. b) For each candidate, the percentage of varieties selected in a subset for field comparison using observed characteristics that were also selected using genomic predicted characteristics. All predictions were run with Random Forest using default settings.

Discussion

Genotypic and phenotypic exploration of the barley prediction panel

The barley prediction panel consisted of 1,086 genotypes with both genetic and phenotypic data. To form the panel, genotypic data from a 50k feature genotyping array were obtained for 1,171 genotypes which covered an estimated ~70% of the active barley 'variety collection' at the time of study. The number of genotypes assembled here forms a powerful resource for applications such as genomic prediction and genome-wide association study (Alqudah et al., 2020; Cericola et al., 2017), highlighting the suitability of the barley prediction panel for the statistical approaches used in this study.

As observed in other barley collections (e.g. Bayer et al., 2017; Cockram et al., 2010; Matthies et al., 2011), there was clear population structure determined by Seasonal type (characteristic no. 90) and Ear: number of rows (no. 39) in the barley prediction panel. Tight clustering was observed for the spring two-row varieties, which is reflective of restricted DUS characteristic space and is in agreement with similar reports on elite spring barley varieties (Yuan et al., 2024; 2025). We found strong correlation between genotypic and phenotypic-based Manhattan distance, with similar values (r = ~ 0.6) found in previous studies (e.g. Jones et al., 2012; Yang et al., 2021). The DUS characteristics investigated here were not all scored across the same number of notes, ranging

between two and nine classes. As Manhattan distances are the sum of absolute differences across variables, these estimates would be biased towards characteristics scored across more notes. Use of min-max characteristic scaling increased correlation between the Manhattan distances, likely due to more equal consideration of binary characteristics that were typically associated with higher heritability estimates (Table 1).

The 712 previously SNP-genotyped barley varieties (Looseley et al., 2020) included in our analysis were previously used to explore DUS characteristic heritability (Yang et al., 2021). Despite the lower number of varieties in their study, the heritability for the DUS characteristics observed by Yang et al. (2021) were very similar to those observed here - indeed, the same DUS characteristics were classified as having low ($h^2 < 0.5$) and high h^2 ($h^2 > 0.5$) heritability, and broadly agree with heritability estimates undertaken in other studies (Cockram et al., 2010). Traits with low heritability are problematic for DUS assessment due to non-genetic factors contributing to more of the observed variation, as has been highlighted in both barley and the closely related cereal crop wheat (Yang et al., 2021; Zanella et al., 2026). Furthermore, the accuracy of genomic prediction and the ability to detect genetic loci are both impacted by heritability (Bernardo et al., 2020). This trend was apparent in our study, whereby DUS characteristics with low heritability were found to have lower prediction accuracies (r), a lower proportion of correctly classified calls (PCCC), and their genetic architecture was rarely found to be associated with a 'major genetic locus' ($-\log_{10}(P) > 10$).

For assessment of distinctness in the current DUS testing pipeline, there is some control for characteristics with low heritability. Firstly, the 'low heritability' group ($h^2 < 0.5$) were typically (but not exclusively) consisted of DUS characteristics scored using a higher number of 'notes' (e.g. scored on a 1-9 scale) and these characteristics are designated wider class widths before assigning distinctness in current testing protocols (Supplementary Table 1). Secondly, assessment of distinctness is only considered in comparison plots within a single field per test year to minimise environmental heterogeneity. However, there are limitations to these considerations (1) characteristic class width did not always scale with heritability. For example, Grain: spiculation of inner lateral nerves (no. 73) and Flag leaf: glaucosity of sheath (no. 15) had very different heritability estimates of 0.75 and 0.14, respectively, and both characteristics shared a class width of three 'notes' in the selection algorithm. (2) Large spatial heterogeneity is common within a field, and it can be difficult to ensure uniformity at the plot level depending on the characteristics under study (Mercer and Hall, 1911). In the current UK DUS assessment protocol for cereal crops only a single plot is used for distinctness assessment per genotype, with the comparison plots grouped together in the same area of the trial to facilitate accurate comparisons. Increasing within-trial replication and using a blocked experimental design could be a useful avenue for improving the accuracy of recording DUS characteristics with low heritability, as is used for many cross-pollinated crops which undergo DUS testing. However, tradeoffs with the cost and increased complexity of the trialing and distinctness assessments would also need to be considered and compared against the necessity and scale of improvement that could be made.

Genomic prediction of barley DUS characteristics with the barley prediction panel

No genomic prediction method is universally optimal for every phenotype and success of a particular approach depends on the architecture of that trait (Daetwyler et al., 2010). Here, heritability appeared the major determinant of prediction accuracy (r) and PCCC, although there were significant differences associated with prediction approach that varied on a characterisitic-by-characteristic basis. However overall, Random Forest (RF) was most effective for traits with higher heritability and fewer characteristic 'notes', although only minor improvements in accuracy were observed compared to the ridge regression-BLUP (rr-BLUP) approach. This is broadly in agreement with previous studies of observed and simulated data in animal species. For example, in the case of binary phenotypes in dairy cattle, RF has been shown to outperform genomic-BLUP in a higher heritability scenario, with a drop in performance under low heritability (Naderi et al., 2016). Similarly, exploration of genomic prediction approaches using simulated livestock data found RF outperformed other approaches when using discrete traits where the phenotype was controlled by a smaller number of genetic loci (González-Recio and Forni, 2011). We found that the genetic architecture of barley DUS characteristics with high heritability were all associated with at least one 'major locus'. Compared to rr-BLUP or RF with default settings, there were small improvements for nearly every characteristic by increasing the probability of selecting and using the most significant genetic markers linked to each of these 'major genetic loci' for splitting in decision trees. Weighting markers linked to genetic loci is an established approach in parametric linear approaches to prediction, at least in animals (Meuwissen et al., 2024; Ren et al., 2020). Fewer studies have explored how RF prediction accuracies could be increased by weighting variables linked to known genetic loci. However, Ghosh and Cabrera (2022) recently proposed a method called 'Enriched Random Forest' that uses weighted random sampling to assign more weight to informative features of the data. Our study provided further evidence that weighting more informative variables, in our case markers linked to 'major genetic loci' (Table 2), can improve RF prediction accuracy.

A more severe drop in RF performance was observed when compared to rr-BLUP for those traits with low heritability and which are phenotyped using relatively high numbers of characteristic 'notes'. rr-BLUP is one of the most commonly used methods of genomic prediction for single traits (Tong and Nikoloski, 2021) and often produces similar levels of accuracy as more complex approaches to prediction (Heslot et al., 2012). For phenotypes controlled by many underlying genes, rr-BLUP has been shown to be an effective approach to prediction (Wang et al., 2015). Significant consideration is needed when using machine learning or artificial intelligence-based approaches to genomic prediction, as potentially small improvements to prediction accuracy or performance observed come with the downside of using a less transparent or more computationally intensive approach (Howard and Lipka, 2025). The RF approach with hyperparameter tuning took several days to complete for each DUS characteristic and advantages or disadvantages observed in terms of r and PCCC were characteristic-dependent.

The requirement to round predicted values to integers would represent a drawback to implementing rr-BLUP in the current DUS testing pipeline, as there was a reduction in prediction accuracy (r) observed for most characteristics with integer rounding. This issue could be circumvented using genomic prediction methods developed for ordinal categorical phenotype data that use a generalized linear mixed models framework (e.g. Montesinos-López et al., 2015). Furthermore, RF prediction accuracies performed worse for ordinal characteristics scored using more 'notes' (those with 1-9 scores) with low heritability. This could be partially attributed to the classification approach not considering the inherent nature of an ordinal characteristic (Janitza et al., 2016).

While the RF prediction generally performed better in terms of PCCC, at lower heritability the rr-BLUP method, which treated each characteristic as numeric, produced better prediction accuracy, indicating rr-BLUP predictions better tracked the variation in the observed data. Ordinal-based RF approaches have been developed (e.g. Hornung, 2019), these tested in conjunction with parametric-based ordinal methods could hold merit for prediction of the DUS characteristics scored with a greater number of 'notes'.

Implications for potential application in DUS testing protocols and next steps towards implementation

Current barley DUS assessment pipelines take up to three years. Low use of DUS characteristic space, as currently observed in spring barley, increases the likelihood of the need for year-3 DUS tests. This in turn increases DUS testing costs and slows down the timescale for genetic improvements in agronomically important traits such as yield and yield quality being realised in farmers' fields. Accurate genomic prediction methodologies could reduce the length of the barley DUS testing cycle by predicting the characteristics of new candidate varieties immediately after being received at the DUS Testing Centre (Figure 4). These predictions would be used to select subsets from the 'variety collection' identified as similar enough to require comparison in field trials candidates in the first year of the DUS pipeline. Roberts et al. (2026) also highlighted that the requirement of a third year of DUS assessment could be removed in crops such as wheat via the implementation of genomic prediction to identify most similar varieties from the variety collection to be grown alongside candidates in the first year of testing. Through this simple change, potential distinctness issues could be identified in year-1 of the process, whereas these are currently first identified in year-2 (Figure 4). These distinctness issues could be explored through detailed examination in the second year of testing via side-by-side comparisons or special testing, and reduce the need for assessment in a third year.

Using default RF settings, ~72% of varieties identified as non-distinct using the observed approach were also found with prediction. With further optimisation of prediction on a trait-by-trait basis, that percentage could be increased. Furthermore, the class widths for each DUS characteristic are currently determined by technical experts based on observed phenotypic variability. Predicted characteristics may vary differently to phenotypic observations and the class widths could be further refined in relation to the observed variability of predicted characteristics.

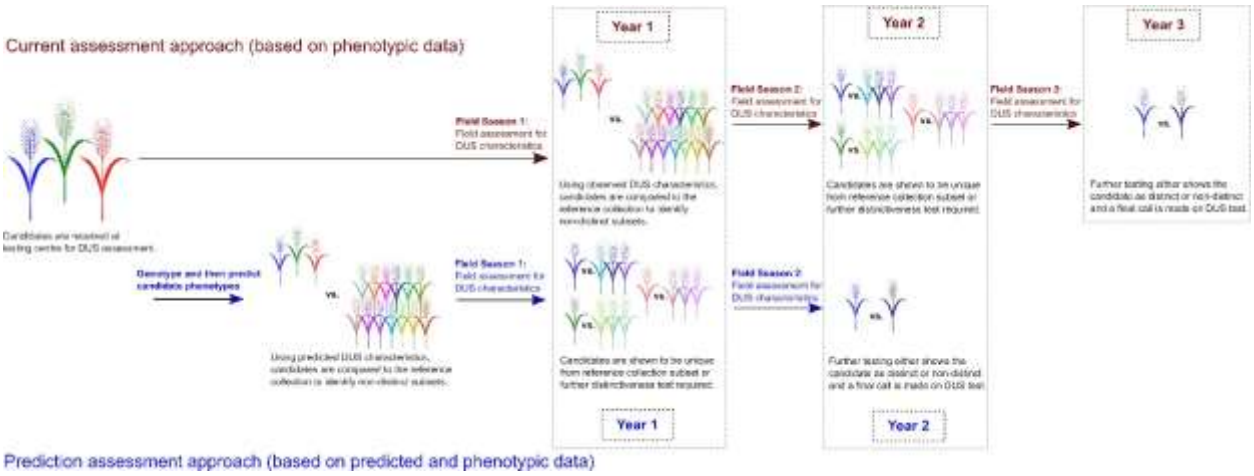


Figure 4 - Schematic showing the current UK barley DUS approach used for distinctness assessment (shown in red), compared with the hypothesised genetic prediction informed approach, which incorporates genomic prediction in year-1 of testing. As distinctness is assessed from year-1 in the prediction informed approach, there is potential for reducing the cycle length by a year.

This work provides proof-of-concept for using genetic prediction in barley DUS testing, the following points are suggested as the next steps towards implementation of this method. (1) Genotyping of the remaining ~30% of the 'variety collection' would allow full coverage of the available genetic diversity. (2) Refinement of the genomic prediction approach on a characteristic-by-characteristic basis, which would involve incorporating 'major genetic loci' weighting, use of optimal hyperparameter combinations and potentially ordinal specific prediction methods for the characteristics with low heritability and a greater number of characteristic 'notes'. (3) Optimisation of DUS distinctness class widths in the selection algorithm to account for prediction characteristic variability. (4) Live parallel field testing of the genetic prediction informed approach alongside the current DUS pipeline, allowing cost-benefit analysis to be determined under real-world settings. (5) Establishing a cost-efficient genotyping solution with adequate turn-around times for real-world implementation. A practical limitation with the proposed approach is that new candidate varieties would need to be quickly genotyped for prediction before selection and inclusion of the relevant comparisons from the variety collection in direct comparisons in year-1. A limitation of array-based genotyping is there is often a minimum sample number requirement to fill a 384-well submission plate, this could lead to cost inefficiency as typically ~45 spring barley candidates are received in the UK annually. Sequence-based approaches such as low-coverage skim-sequencing following by imputation have no minimum sample number requirement and could prove a more cost-effective solution (Sthapit et al., 2025).

Final conclusions

Molecular markers are routinely used in crop breeding and research to accelerate germplasm assessment and development timelines. Phenotypic observations remains at the core of barley DUS assessment, but these pipelines could be further optimised via use of genomic information. Here, we suggested an approach for the inclusion of genetic markers that does not replace traditional morphological assessment; rather, it offers a potential route to shortening the length of current

DUS assessment methodology which could improve the responsiveness of variety development and release.

Acknowledgements

We wish to thank Adrian Roberts and Daniel Tolhurst for the interesting scientific discussions relating to this work, Bethany Love for preliminary quality checks on the genetic marker data, and the feedback of colleagues at Niab. We thank the Defra Plants, Varieties and Seeds team, Katherine Deeks, Sigurd Ramans-Harborough, Christina Bradley and Lucia Micheluzzi for commissioning the project and providing support and feedback across the duration of the work.

Author Contributions

JC and **MDW** conceptualised this study. **VEM** and **RH** prepared samples and extracted DNA for genotyping, and **VEM** complied the phenotypic data. **MM**, **LR** and **PDS** provided further genomic data and conducted the genetic marker calling. Quality control and data analysis were completed by **TICW**. **NFG** contributed guidance with machine-learning based genomic prediction. The first draft of the manuscript was prepared by **TICW** and **VEM**, with all authors contributing to further improvements of the manuscript.

Funding

This work was funded by the Department for Environment, Food and Rural Affairs (Defra) Plants, Varieties and Seeds team, under the Research, Development and Evidence (RDE) framework.

Data availability

The analysis scripts used in this study are available from the GitHub repository: www.github.com/TallyWright/DUS-barley-GP. Variety collection data collected during DUS testing is confidential and not available in the public domain, with the exception of publicly available variety descriptions which can be found here: https://www.niab.com/services/seed-certification/botanical-descriptions-varieties. The original 50k genetic marker data for the IMPROMALT genotypes are available from the James Hutton Institute on request with approval of the IMPROMALT consortium.

References

Alqudah AM, Sallam A, Baenziger PS, Börner A (2020) GWAS: Fast-forwarding gene identification and characterization in temperate cereals: lessons from barley – A review. Journal of Advanced Research 22: 119–135. https://doi.org/10.1016/j.jare.2019.10.013

Animal and Plant Health Agency (2014) DUS protocols for testing plant varieties. GOV.UK. https://www.gov.uk/guidance/dus-protocols-for-testing-plant-varieties

Arrieta M, Macaulay M, Colas I, Schreiber M, Shaw PD, Waugh R and Ramsay L (2021) An induced mutation in *HvRECQL4* increases the overall recombination and restores fertility in a barley *HvMLH3* mutant background. Frontiers in Plant Science 12: 706560. https://doi.org/10.3389/fpls.2021.706560

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67: 1–48. https://doi.org/10.18637/jss.v067.i01

Bayer MM, Rapazote-Flores P, Ganal M, Hedley PE, Macaulay M, Plieske J, Ramsay L, Russell J, Shaw PD, Thomas W, Waugh R (2017) Development and evaluation of a barley 50k iSelect SNP array. Frontiers in Plant Science 8. https://doi.org/10.3389/fpls.2017.01792

Bernardo R (2020) Breeding for quantitative traits in plants, 3rd ed. Stemma Press, Woodbury, MN.

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. Crop Science 47: 1082–1090. https://doi.org/10.2135/cropsci2006.11.0690

Bull H, Casao MC, Zwirek M, Flavell AJ, Thomas WTB, Guo W, Zhang R, Rapazote-Flores P, Kyriakidis S, Russell J, Druka A, McKim SM, Waugh R (2017) Barley *SIX-ROWED SPIKE3* encodes a putative Jumonji C-type H3K9me2/me3 demethylase that represses lateral spikelet fertility. Nature Communications 8. https://doi.org/10.1038/s41467-017-00940-7

Cericola F, Jahoor A, Orabi J, Andersen JR, Janss LL, Jensen J (2017) Optimizing training population size and genotyping strategy for genomic prediction using association study results and pedigree information. A case of study in advanced wheat breeding lines. PLOS ONE 12: e0169606. https://doi.org/10.1371/journal.pone.0169606

Chen X, Ishwaran H (2012) Random forests for genomic data analysis. Genomics 99: 323–329. https://doi.org/10.1016/j.ygeno.2012.04.003

Cockram J, Jones H, Norris C, O'Sullivan DM (2012) Evaluation of diagnostic molecular markers for DUS phenotypic assessment in the cereal crop, barley (*Hordeum vulgare* ssp. *vulgare* L.). Theoretical and Applied Genetics 125: 1735–1749. https://doi.org/10.1007/s00122-012-1950-3

Cockram J, White J, Zuluaga DL, Smith D, Comadran J, Macaulay M, Luo Z, Kearsey MJ, Werner P, Harrap D, Tapsell C, Liu H, Hedley PE, Stein N, Schulte D, Steuernagel B, Marshall DF, Thomas WTB, Ramsay L, Mackay I, Balding DJ, Waugh R, O'Sullivan DM, Booer C, Pike S, Hamilton G, Jellis G, Davies N, Ross A, Bury P, Habgood R, Klose S, Vequaud D, Christerson T, Brosnan J, Newton A, Russell J, Shaw P, Bayles R, Wang M (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proceedings of the National Academy of Sciences 107: 21611–21616. https://doi.org/10.1073/pnas.1010179107

Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA (2010) The impact of genetic architecture on genome-wide evaluation methods. Genetics 185: 1021–1031. https://doi.org/10.1534/genetics.110.116855

De la Fuente Cantó C, Russell J, Hackett CA, Booth A, Dancey S, George TS, Waugh R (2018) Genetic dissection of quantitative and qualitative traits using a minimum set of barley recombinant chromosome substitution lines. BMC Plant Biology 18. https://doi.org/10.1186/s12870-018-1527-7

Distelfeld A, Li C, Dubcovsky J (2009) Regulation of flowering in temperate cereals. Current Opinion in Plant Biology 12: 178–184. https://doi.org/10.1016/j.pbi.2008.12.010

Dubcovsky J, Chen C, Yan L (2005) Molecular characterization of the allelic variation at the *VRN-H2* vernalization locus in barley. Molecular Breeding 15: 395–407. https://doi.org/10.1007/s11032-005-0084-6

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4: 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Faure S, Higgins J, Turner A, Laurie DA (2007) The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*). Genetics 176: 599–609. https://doi.org/10.1534/genetics.106.069500

Faure S, Turner AS, Gruszka D, Christodoulou V, Davis SJ, von Korff M, Laurie DA (2012) Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. Proceedings of the National Academy of Sciences 109: 8328–8333. https://doi.org/10.1073/pnas.1120496109

Fradgley N, Gardner KA, Bentley AR, Howell P, Mackay IJ, Scott MF, Mott R, Cockram J (2023) Multi-trait ensemble genomic prediction and simulations of recurrent selection highlight importance of complex trait genetic architecture for long-term genetic gains in wheat. in Silico Plants 5: 1–20. https://doi.org/10.1093/insilicoplants/diad002

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Molecular Biology Reporter 13: 207–209. https://doi.org/10.1007/BF02670897

Ghosh D, Cabrera J (2022) Enriched random forest for high dimensional genomic data. IEEE/ACM Transactions on Computational Biology and Bioinformatics 19: 2817–2828. https://doi.org/10.1109/tcbb.2021.3089417

Covarrubias-Pazaran G (2016) Genome assisted prediction of quantitative traits using the R package sommer. PLoS ONE 11: 1–15. https://doi.org/10.1371/journal.pone.0156744

González-Recio O, Forni S (2011) Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. Genetics Selection Evolution 43. https://doi.org/10.1186/1297-9686-43-7

Hasan N, Choudhary S, Naaz N, Sharma N, Laskar RA (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. Journal of Genetic Engineering and Biotechnology 19: 128. https://doi.org/10.1186/s43141-021-00231-1

Heslot N, Yang H, Sorrells ME, Jannink J (2012) Genomic selection in plant breeding: A comparison of models. Crop Science 52: 146–160. https://doi.org/10.2135/cropsci2011.06.0297

Hickey JM, Chiurugwi T, Mackay I, Powell W (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nature Genetics 49: 1297–1303. https://doi.org/10.1038/ng.3920

Hornung R (2019) Ordinal forests. Journal of Classification 37: 4–17. https://doi.org/10.1007/s00357-018-9302-x

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biometrical Journal 50: 346–363. https://doi.org/10.1002/bimj.200810425

Howard R, Lipka AE (2025) Genomic selection and reproducibility: Are complex models distracting us from true scientific validity in the presence of genotype-by-environment interaction? G3: Genes, Genomes, Genetics 15. https://doi.org/10.1093/g3journal/jkaf244

International Union for the Protection of New Varieties of Plants (UPOV) (2025) Barley: Guidelines for the conduct of tests for distinctness, uniformity and stability (Test Guidelines No. TG/19/11 Corr.). UPOV, Geneva.

International Union for the Protection of New Varieties of Plants (UPOV) (2020) Guidance on the use of biochemical and molecular markers in the examination of distinctness, uniformity and stability (Technical Report No. TGP/15/3). UPOV, Geneva.

International Union for the Protection of New Varieties of Plants (UPOV) (2023) Kale: Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability (Test Guidelines No. TG/90/7). UPOV, Geneva.

International Union for the Protection of New Varieties of Plants (UPOV) (2008) TGP/4: Constitution and maintenance of variety collections: Associated document to the general introduction to the examination of DUS and the development of harmonized descriptions of new varieties of plants (TG/1/3) (Technical Report No. TGP/4/1). UPOV, Geneva.

International Union for the Protection of New Varieties of Plants (UPOV) (2024) Tomato: Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability (Test Guidelines No. TG/44/12). UPOV, Geneva.

Ishwaran H, Kogalur UB (2025) Fast unified random forests for survival, regression, and classification (RF-SRC). R package version 3.4.1. https://cran.r-project.org/package=randomForestSRC.

Jamali SH, Cockram J, Hickey LT (2020) Is plant variety registration keeping pace with speed breeding techniques? Euphytica 216: 131. https://doi.org/10.1007/s10681-020-02666-y

Janitza S, Tutz G, Boulesteix A-L (2016) Random forest for ordinal responses: Prediction and variable selection. Computational Statistics & Data Analysis 96: 57–73. https://doi.org/10.1016/j.csda.2015.10.005

Jayakodi M, Lu Q, Pidon H, Rabanus-Wallace MT, Bayer M, Lux T, Guo Y, Jaegle B, Badea A, Bekele W, Brar GS, Braune K, Bunk B, Chalmers KJ, Chapman B, Jørgensen ME, Feng J-W, Feser M, Fiebig A, Gundlach H, Guo W, Haberer G, Hansson M, Himmelbach A, Hoffie I, Hoffie RE, Hu H, Isobe S, König P, Kale SM, Kamal N, Keeble-Gagnère G, Keller B, Knauft M, Koppolu

R, Krattinger SG, Kumlehn J, Langridge P, Li C, Marone MP, Maurer A, Mayer KFX, Melzer M, Muehlbauer GJ, Murozuka E, Padmarasu S, Perovic D, Pillen K, Pin PA, Pozniak CJ, Ramsay L, Pedas PR, Rutten T, Sakuma S, Sato K, Schüler D, Schmutzer T, Scholz U, Schreiber M, Shirasawa K, Simpson C, Skadhauge B, Spannagl M, Steffenson BJ, Thomsen HC, Tibbits JF, Nielsen MTS, Trautewig C, Vequaud D, Voss C, Wang P, Waugh R, Westcott S, Rasmussen MW, Zhang R, Zhang X-Q, Wicker T, Dockter C, Mascher M, Stein N (2024) Structural variation in the pangenome of wild and domesticated barley. Nature 636: 654–662. https://doi.org/10.1038/s41586-024-08187-1

Jones H, Norris C, Cockram J, Lee D (2013) Variety Protection and Plant Breeders' Rights in the 'DNA era', in: Diagnostics in Plant Breeding. Springer, Netherlands, pp. 396–402.

Jones H, Norris C, Smith D, Cockram J, Lee D, O'Sullivan DM, Mackay I (2012) Evaluation of the use of high-density SNP genotyping to implement UPOV model 2 for DUS testing in barley. Theoretical and Applied Genetics 126: 901–911. https://doi.org/10.1007/s00122-012-2024-2

Kim J-S, Takahagi K, Inoue K, Shimizu M, Uehara-Yamaguchi Y, Kanatani A, Saisho D, Nishii R, Lipka AE, Hirayama T, Sato K, Mochida K (2023) Exome-wide variation in a diverse barley panel reveals genetic associations with ten agronomic traits in eastern landraces. Journal of Genetics and Genomics 50: 241–252. https://doi.org/10.1016/j.jgg.2022.12.001

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proceedings of the National Academy of Sciences 104: 1424-1429. https://doi.org/10.1073/pnas.0608580104

Kuhn M (2008) Building predictive models in r using the caret package. Journal of Statistical Software 28: 1–26. https://doi.org/10.18637/jss.v028.i05

Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software 82: 1–26. https://doi.org/10.18637/jss.v082.i13

Lange TM, Gültas M, Schmitt AO, Heinrich F (2025) optRF: Optimising random forest stability by determining the optimal number of trees. BMC Bioinformatics 26. https://doi.org/10.1186/s12859-025-06097-1

Lenth RV (2024) emmeans: Estimated marginal means, aka least-squares means. R package version 1.10.4. https://cran.r-project.org/package=emmeans

Looseley ME, Ramsay L, Bull H, Swanston JS, Shaw PD, Macaulay M, Booth A, Russell JR, Waugh R, the IMPROMALT Consortium, Thomas WTB (2020) Association mapping of malting quality traits in UK spring and winter barley cultivar collections. Theoretical and Applied Genetics 133: 2567–2582. https://doi.org/10.1007/s00122-020-03618-9

Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, Bayer M, Ramsay L, Liu H, Haberer G, Zhang X-Q, Zhang Q, Barrero RA, Li L, Taudien S, Groth M, Felder M, Hastie A, Šimková H, Staňková H, Vrána J, Chan S, Muñoz-Amatriaín M, Ounit R, Wanamaker S, Bolser D, Colmsee C, Schmutzer T, Aliyeva-

Schnorr L, Grasso S, Tanskanen J, Chailyan A, Sampath D, Heavens D, Clissold L, Cao S, Chapman B, Dai F, Han Y, Li H, Li X, Lin C, McCooke JK, Tan C, Wang P, Wang S, Yin S, Zhou G, Poland JA, Bellgard MI, Borisjuk L, Houben A, Doležel J, Ayling S, Lonardi S, Kersey P, Langridge P, Muehlbauer GJ, Clark MD, Caccamo M, Schulman AH, Mayer KFX, Platzer M, Close TJ, Scholz U, Hansson M, Zhang G, Braumann I, Spannagl M, Li C, Waugh R, Stein N (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544: 427–433. https://doi.org/10.1038/nature22043

Matthies IE, van Hintum T, Weise S, Röder MS (2011) Population structure revealed by different marker types (SSR or DArT) has an impact on the results of genome-wide association mapping in European barley cultivars. Molecular Breeding 30: 951–966. https://doi.org/10.1007/s11032-011-9678-3

Mercer WB, Hall AD (1911) The experimental error of field trials. The Journal of Agricultural Science 4: 107–132. https://doi.org/10.1017/s002185960000160x

Meuwissen T, Eikje LS, Gjuvsland AB (2024) GWABLUP: Genome-wide association assisted best linear unbiased prediction of genetic values. Genetics Selection Evolution 56. https://doi.org/10.1186/s12711-024-00881-y

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Montesinos López OA, Montesinos López A, Crossa J (2022) Random forest for genomic prediction, in: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer International Publishing, Cham, pp. 633–681. https://doi.org/10.1007/978-3-030-89010-0_15

Montesinos-López OA, Montesinos-López A, Pérez-Rodríguez P, de los Campos G, Eskridge K, Crossa J (2015) Threshold models for genome-enabled prediction of ordinal categorical traits in plant breeding. G3 Genes|Genomes|Genetics 5: 291–300. https://doi.org/10.1534/g3.114.016188

Naderi S, Yin T, König S (2016) Random forest estimation of genomic breeding values for disease susceptibility over different disease incidences and genomic architectures in simulated cow calibration groups. Journal of Dairy Science 99: 7261–7273. https://doi.org/10.3168/jds.2016-10887

Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Borman T, Carvalho G, Chirico M, De Cáceres M, Durand S, Evangelista HBA, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill MO, Lahti L, Martino C, McGlinn D, Ouellette M-H, Ribeiro Cunha E, Smith T, Stier A, Ter Braak CJF, Weedon J (2025) Vegan: Community ecology package. R package version 2.7-1. https://cran.r-project.org/web/packages/vegan/index.html

Paradis E, Schliep K (2019) Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35: 526–528. https://doi.org/10.1093/bioinformatics/bty633

Percival-Alwyn L, Barnes I, Clark MD, Cockram J, Coffey MP, Jones S, Kersey PJ, Kidner CA, Kosiol C, Li B, Marsh WA, Zhou J, Caccamo M, Milne I (2024) UKCropDiversity-HPC: A

collaborative high-performance computing resource approach for sustainable agriculture and biodiversity conservation. Plants People Planet 7: 969–977. https://doi.org/10.1002/ppp3.10607

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N, Hayes PM, Lundqvist U, Franckowiak JD, Close TJ, Muehlbauer GJ, Waugh R (2011) *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. Nature Genetics 43: 169-172. https://doi.org/10.1038/ng.745

Ren D, An L, Li B, Qiao L, Liu W (2020) Efficient weighting methods for genomic best linear-unbiased prediction (BLUP) adapted to the genetic architectures of quantitative traits. Heredity 126: 320–334. https://doi.org/10.1038/s41437-020-00372-y

Roberts AMI, Neugebauer K, Ewaoluwagbemiga EO, Milbourne D, Byrne S, Cockram J, Zanella CM, Wallace M, Niedermayer M, Andreani L, Pecs M, van der Kooij WWP, Schmid K (2026) Integrating genomic prediction into crop DUS testing: new frameworks for reference collection management and distinctness assessment. Theoretical and Applied Genetics 139. https://doi.org/10.1007/s00122-026-05198-6

Rosyara UR, De Jong WS, Douches DS, Endelman JB (2016) Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome 9. https://doi.org/10.3835/plantgenome2015.08.0073

Stekhoven DJ, Buehlmann P (2012) MissForest - non-parametric missing value imputation for mixed-type data. Bioinformatics 28: 112–118. https://doi.org/10.1093/bioinformatics/btr597

Sthapit SR, Crain J, Larson S, Anderson JA, Bajgain P, DeHaan LR, Poland J (2025) A low-coverage skim-sequencing and imputation pipeline for genomic selection. The Plant Genome 18. https://doi.org/10.1002/tpg2.70139

Strygina KV, Börner A, Khlestkina EK (2017) Identification and characterization of regulatory network components for anthocyanin synthesis in barley aleurone. BMC Plant Biology 17. https://doi.org/10.1186/s12870-017-1122-3

Tong H, Nikoloski Z (2021) Machine learning approaches for crop improvement: Leveraging phenotypic and genotypic big data. Journal of Plant Physiology 257: 153354. https://doi.org/10.1016/j.jplph.2020.153354

VanRaden PM (2008) Efficient methods to compute genomic predictions. Journal of Dairy Science 91: 4414–4423. https://doi.org/10.3168/jds.2007-0980

Voss-Fels KP, Cooper M, Hayes BJ (2018) Accelerating crop genetic gains with genomic selection. Theoretical and Applied Genetics 132: 669–686. https://doi.org/10.1007/s00122-018-3270-8

Voss-Fels KP, Stahl A, Hickey LT (2019) Q&A: modern crop breeding for future food security. BMC Biology 17. https://doi.org/10.1186/s12915-019-0638-4

Wang X, Yang Z, Xu C (2015) A comparison of genomic selection methods for breeding value prediction. Science Bulletin 60: 925–935. https://doi.org/10.1007/s11434-015-0791-2

Whittaker JC, Thompson R, Denham MC (2000) Marker-assisted selection using ridge regression. Genetical Research 75: 249–252. https://doi.org/10.1017/s0016672399004462

Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Wright TIC, Horsnell R, Love B, Burridge AJ, Gardner KA, Jackson R, Leigh FJ, Ligeza A, Heuer S, Bentley AR, Howell P (2024) A new winter wheat genetic resource harbors untapped diversity from synthetic hexaploid wheat. Theoretical and Applied Genetics 137. https://doi.org/10.1007/s00122-024-04577-1

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-1644. https://doi.org/10.1126/science.1094305

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. Proceedings of the National Academy of Sciences USA 103: 19581-19586. https://doi.org/10.1073/pnas.0607142103

Yang CJ, Russell J, Ramsay L, Thomas W, Powell W, Mackay I (2021) Overcoming barriers to the registration of new plant varieties under the DUS system. Communications Biology 4: 302. https://doi.org/10.1038/s42003-021-01840-9

Yuan Z, Rembe M, Mascher M, Stein N, Himmelbach A, Jayakodi M, Börner A, Oldach K, Jahoor A, Jensen JD, Rudloff J, Dohrendorf V.-E, Kuhfus LP, Dyrszka E, Conte M, Hinz F, Trouchaud S, Reif JC, El Hanafi S (2025) High-quality phenotypic and genotypic dataset of barley genebank core collection to unlock untapped genetic diversity. GigaScience 14. https://doi.org/10.1093/gigascience/giae121

Yuan Z, Rembe M, Mascher M, Stein N, Jayakodi M, Börner A, Oldach K, Jahoor A, Jensen JD, Rudloff J, Dohrendorf V-E, Kuhfus LP, Dyrszka E, Conte M, Hinz F, Trouchaud S, Reif JC, El Hanafi S (2024) Capitalizing on genebank core collections for rare and novel disease resistance loci to enhance barley resilience. Journal of Experimental Botany 75: 5940–5954. https://doi.org/10.1093/jxb/erae283

Zanella CM, Horsnell R, Love B, Wright TIC, Taferner-Kriegl J, Flamm C, Andreani L, Delogu C, McMillan V, Wallace M, Scott E, Cockram J (2026) Genome-wide association study (GWAS) identifies genetic loci controlling Distinctness, Uniformity, and Stability (DUS) traits in wheat. Theoretical and Applied Genetics 139. https://doi.org/10.1007/s00122-025-05130-4

Zhang Z, Ober U, Erbe M, Zhang H, Gao N, He J, Li J, Simianer H (2014) Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. PLoS ONE 9: e93017. https://doi.org/10.1371/journal.pone.0093017

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryText.docx
- SupplementaryTables.xlsx
- SupplementaryFigures.pptx