TREND, POPULATION STRUCTURE AND TRAIT MAPPING FROM 15 YEARS OF NATIONAL VARIETAL TRIALS OF UK WINTER WHEAT.

Oluwaseyi Shorinola [*†], **James Simmonds** [*], **Luzie U. Wingen** [*], **Keith Gardner** [‡], **Cristobal Uauy** [*]

[*] John Innes Centre, Norwich Research Park, Norwich UK, NR4 7UH.

[†]Bioscience eastern and central Africa – International Livestock Research Institute (BecA-ILRI), Nairobi, 00100, Kenya.

[‡]National Institute of Agricultural Botany, 93 Lawrence Weaver Road, Cambridge UK, CB3 0LE.

Running Title: Fifteen years of UK Wheat breeding.

Corresponding Author: Oluwaseyi Shorinola, John Innes Centre, Norwich Research Park, Norwich UK, NR4 7UH, Oluwaseyi.shorinola@jic.ac.uk, o.shorinola@cgiar.org.

Keywords: Wheat, SNP, GWAS, *NAM-A1*, Trend Analysis.

ABSTRACT

There are now a rich variety of genomic and genotypic resources available to wheat researchers and breeders. However, the generation of high-quality and field-relevant phenotyping data which is required to capture the complexities of gene x environment interactions remains a major bottleneck. Historical datasets from national variety performance trials (NVPT) provide sufficient dimensions, in terms of numbers of years and locations, to examine phenotypic trends and study gene x environment interactions. Using NVPT for winter wheat varieties grown in the UK between 2002 – 2017, we examined temporal trends for eight traits related to yield, adaptation, and grain quality performance. We show a non-stationary linear trend for yield, grain protein content, HFN and days to ripening. Our data also show high environmental stability for yield, grain protein content and specific weight in UK winter wheat varieties and high environmental sensitivity for Hagberg Falling Number. Using the historical NVPT data in a genome-wide association analysis, we uncovered a significant marker-trait association peak on wheat chromosome 6A spanning the *NAM-A1* gene that have been previously associated with early senescence. Together our results show the value of utilizing the data routinely collected during variety evaluation process for examining breeding progress and the genetic architecture of important traits.

INTRODUCTION

Over the last three years, there has been a rapid surge in the development of genomic resources for wheat (reviewed in Adamski *et al.* 2020). This includes a chromosome-scale reference assembly of the Chinese Spring cultivar (RefSeqv1) and a pan-genome resource comprised of chromosome and scaffold-level assemblies of 15 hexaploid wheat cultivars (IWGSC *et al.* 2018; Walkowiak *et al.* 2020). There is also a wide range of array-based (Axiom-35K, iSelect 90K, Axiom-660K and Axiom-820K; Wang *et al.* 2014; Winfield *et al.* 2016; Allen *et al.* 2017), sequencing-based (e.g DARTSeq, RADSeq) or PCR-based (e.g KASP, TaqMan, rhAmp; Semagn *et al.* 2014; Ayalew *et al.* 2019) SNP genotyping assays available to wheat researchers and breeders. There have also been efforts to re-sequence different wheat populations either through reduced-representation sequencing approach like exome-capture and sequencing (e.g He *et al.* 2019; Krasileva *et al.* 2017; Jordan *et al.* 2015) or through whole genome resequencing (e.g Cheng *et al.* 2019; Scott *et al.* 2020). This preponderance of genomics and genotypic data which are available in open-access repositories (e.g. EnsemblPlants, CerealsDB; Bolser *et al.* 2016; Howe *et al.* 2020; Wilkinson *et al.* 2020) now makes it possible to map traits at high-resolution (e.g Walkowiak *et al.* 2020), examine population diversity at whole genome levels or in breeding units (haplotypes: e.g Brinton *et al.* 2020; Scott *et al.* 2020), and implement genome-assisted breeding schemes using marker-assisted and/or genomic selection (e.g Sweeney *et al.* 2019; Rasheed and Xia 2019).

Despite these advances, the generation of high-quality and field-relevant phenotyping data remains a major bottleneck. Modern phenomics platforms have improved

78 phenotyping throughput and precision under controlled conditions, but these do not
79 always capture the environmental effects experienced under real-world farming
80 conditions (Yang *et al.* 2020). Given climate change projections of fluctuating radiation,
81 heat and precipitation patterns in major wheat growing areas (including the UK),
82 breeding for phenotypic stability and understanding complex gene x environment
83 interactions is of high priority (Semenov 2009; Trnka *et al.* 2019).

84 Due to their large scale and multi-environment (years and locations) design, historical
85 dataset from national variety performance trials (NVPT) provide sufficient dimensions, in
86 terms of years and locations to examine phenotypic trends and study gene x
87 environment interactions. These historical datasets are, however, incomplete by design
88 because of, for example, changes in the number and specific set of varieties trialed and
89 changes in the field sites used from year to year. Previous studies have analyzed NPVT
90 data for wheat in the UK (Silvey 1981; Mackay *et al.* 2011) and similar analyses of
91 historical data have been conducted elsewhere (e.g Crossa *et al.* 2007; Pozniak *et al.*
92 2012).

93 In the UK, new wheat varieties undergo statutory tests before they are registered on the
94 National List (NL). Registered varieties are subsequently introduced (or maintained on)
95 the UK Recommended List (RL) after undergoing independent non-statutory NPVT
96 managed by the Agriculture and Horticulture Development Board (AHDB, formerly
97 Home-Grown Cereals Authority). The NL serves as variety registry while the RL is used
98 as a reference by farmers for variety selection. Mackay et al. (2011) re-analyzed data
99 from the UK NL and RL trials conducted between 1948 – 2007, and found significant
100 yield improvement that was mostly attributed to plant breeding. In the present study, we

analyzed data from the UK RL NVPT for winter wheat between 2002 - 2017 and used this to examine temporal trends in eight yield, adaptation, and grain quality traits. We also demonstrate the usefulness of these NVPT dataset for trait mapping to uncover loci of breeding importance.

MATERIALS AND METHODS

NVPT datasets

We downloaded result files for the NVPT for winter wheat in the UK from 2002 – 2010 and 2012 - 2017 from the AHDB website (accessible at: https://ahdb.org.uk/knowledge-library/recommended-lists-for-cereals-and-oilseeds-rl-harvest-results-archive). We focused our study on data for eight traits including yield, adaptation and grain quality traits. Yield and height data were collected from treated and untreated trials. The treated trials included management for diseases (fungicide spray) while the untreated trials did not include disease management. Both trials were managed under standard husbandry practices including the application of plant growth regulator (PGR), herbicide, fertiliser and pest control management as recommended by AHDB. Details of the AHDB RL trial protocol is accessible at: https://ahdb.org.uk/rlprotocols. Before analyses, we filtered the dataset to remove observations with unknown locations or from locations where trials were abandoned. Varieties that were trialed in a single year were also omitted. The nomenclature of varieties, locations and counties were standardized in cases

where different designation or acronyms were used for the same variety, location or county across different years. After filtering, the distribution of the observations obtained for each of the eight target traits resemble a bell curve suggesting normal distribution (Figure S1).

Germplasm

Data for a total of 168 varieties were used in this study. These include 133 varieties whose phenotype information were obtained from the AHDB website as described above. For 139 varieties, which included additional 35 pre-2002 UK wheat varieties, genotype data from the Axiom-35K array was used as described below. The number of varieties used for each analysis in this study are detailed in Figure S2.

Statistical Analyses

We used a two-stage approach to examine the linear trend of trait from the NVPT data. First, we fitted a linear mixed model (LMM) to the NVPT data using restricted maximum likelihood (REML) estimation. The model was implemented using the lme4 package in R as:

$$Y_{ijk} = \mu + v_i + y_j + s_{jk} + e_{ijk}$$

Y_{ijk} is the historical performance of variety i in year j at location k. μ is the overall mean performance of all varieties, v_i is the effect of variety i, y_j is the effect of year j (the calendar year of the trial) and S_{jk} is the effect of location k within year j. e_{ijk} is the residual variance arising from factors not accounted for in the model including variety x year interaction. As our main interest was the performance for each variety, the variety effect was fitted as fixed factor while the year and site (nested within year) were fitted as

144 random factors. This is slightly different to the strategy used by Mackay et al. (2011),
145 which also included calendar year as a fixed factor to account for the long year interval
146 (1948 – 2002) examined and changes in trial management system across these years.
147 Given the short interval examined in this study, we believe the management systems
148 were fairly uniform across the trial year. We derived estimates for the varieties means
149 (EVM) from the LMM. Second, we used a linear model to regress the EVM derived from
150 the LMM above against the year the variety was first entered into the NVPT. For trait
151 comparison between end-use groups, Analysis of Variance (ANOVA) followed by post-
152 hoc TukeyHSD was used to evaluate and compare significant difference in EVM of
153 varieties belonging to different end-use groups. The lstrend function implemented in the
154 *R* lsmeans package (Lenth 2016) was used to estimate and compare slopes of the
155 linear regression between groups. For slope comparisons between the four end-use
156 groups, the adjusted P value is presented based on Tukey's method of comparison.

157 We used the Finlay Wilkinson regression to examine phenotype stability (Finlay and
158 Wilkinson 1963). The original Finlay Wilkinson regression used by breeders to examine
159 varietal adaptability is not best suited for data from incomplete trial design as the
160 environment means used for normalizing varietal performance are biased due to
161 incomplete replication of varieties across all environments. To circumvent this bias in
162 our analysis, we used the Bayesian method proposed by Su et al., (2006) and
163 implemented in the *R* package FW (Lian and de los Campos 2016). Only varieties that
164 were trialed in more than three years were used for this analysis. The mean values for
165 each variety in each year were used as input. The model was fitted with the Bayesian
166 "gibbs" method, with 50000 iterations and 5000 burnIn rate as suggested for wheat

167 analyses in the FW package paper (Lian and de los Campos 2016). The FW
168 coefficients are presented as b + 1 which describes expected change in variety
169 performance per unit change of the environment effect (Lian and de los Campos 2016).

170 **Genotyping, population structure and association analysis**

171 A subset of 139 modern varieties and historic cultivars were genotyped using the
172 Axiom-35K array (Allen *et al.* 2017). We filtered the genotype data to include only sites
173 with > 0.05 minor allele frequencies. Marker with heterozygous calls but that were
174 missing one of the homozygous calls (e.g markers with AA and AC but missing CC)
175 were also removed as these are likely due to wrong genotype assignment during
176 automated genotype cluster analysis. The markers were filtered to remove pair of loci
177 with high linkage disequilibrium (R^2 > 0.75). This was done to remove biases arising
178 from high LD loci (such as from introgression from wild relatives) that can bias the
179 contributions of such loci in population structure analysis. To assign physical positions
180 to the Axiom markers, their sequences were used as queries in BLASTn alignments
181 against the IWGSC RefSeqv1.0 assembly (IWGSC *et al.* 2018) as described in Brinton
182 et al. (2020) and the best hits on each of the three wheat homoeologous genomes (A, B
183 and D) were recorded. Of these, the correct homoeologous chromosome was selected
184 using genetic mapping information from 13 populations (Gardiner *et al.* 2019) where
185 available for each marker. Otherwise, the highest BLASTn score was used to select the
186 homoeologous chromosome. In case of conflicting genetic mapping results for the
187 correct chromosome between the mapping populations, the most frequent outcome was
188 used.

Population structure analysis was done using discriminant analysis of principal component (DAPC) as implemented in the Adegenet *R* package (Jombart and Ahmed 2011). For this, the number of population cluster (*k*) was determined by kmeans clustering using a range of *k*. The *k* with the minimum Bayesian Information Criterion was selected as the optimum *k*. To increase the accuracy of grouping, 50 iterations of the kmeans clustering algorithm was run and the population group to which a variety was most frequently assigned was selected. Also, the cross-validation function (xvalDapc) was used to select the optimum number of principal components to use for DAPC.

GWASpoly – a *R* package for association analysis in polyploid crop, was used for GWAS (Rosyara *et al.* 2016). We used a K+Q mixed model where K represents the kinship matrix describing the relatedness between the varieties and Q represents the population grouping derived from the DAPC analysis. A Bonferroni threshold with adjusted *P* value below 0.05 was used to select markers with significant association with the trait of interest.

Data Availability

The original data files for the trials described in this study can be downloaded from the AHDB website at: https://ahdb.org.uk/knowledge-library/recommended-lists-for-cereals-and-oilseeds-rl-harvest-results-archive. As data for different traits are combined in these original files, we re-organized the files to separate the data for each trait into separate files. The re-organized files are available at Zenodo: https://doi.org/10.5281/zenodo.4761528. The QC-filtered trial data used for subsequent

211 analyses are presented in Table S1. Table S2 contains the end-use group information
212 and linear-mixed-model-derived EVM for the varieties trialed. Table S3 contains the FW
213 coefficients for each variety used in the FW regression analysis. The filtered Axiom-35K
214 genotyping data and their genome distribution are presented in Table S4 and Table S5,
215 respectively. Table S6 contains the population group information for each variety
216 genotyped.

217

218

219 # RESULTS

220 **Estimates from multi-environment trial capture expected relationship between**
221 **traits**

222 We analyzed the historical data set of the UK RL NVPT from 2002 to 2017. We focused
223 our analyses on six traits of agronomic and economic importance: yield, plant height,
224 days to ripening, Hagberg Falling Number (HFN), grain protein content and specific
225 weight. For yield and plant height, we analyzed data coming from (fungicide) treated
226 and untreated trials. This results in a final dataset for eight traits. After quality controls
227 (described in Materials and Methods), we retained 52,152 observations for these eight
228 traits from 133 winter wheat varieties (Table S1). These 133 varieties were phenotyped
229 in at least two years across a combined 162 locations, with a subset of 95 locations
230 being used for evaluations in two or more years. Table 1 details the number of varieties

231 phenotyped for each trait and the number of locations and year-location combinations
232 used. The trial locations were spread across 43 counties and unitary authorities in
233 England, Wales, Scotland, and Northern Ireland as shown in Figure 1.

234 Using a linear mixed model that accounted for variation arising from the different years
235 and trial locations, we derived estimates for variety mean (hereafter referred to as EVM)
236 for each variety for each trait (Table S2). Correlation analysis using the EVM captured
237 expected patterns of relationship between the measured traits (Figure 2). We observed
238 significant positive correlations between treated and untreated trials for height and yield,
239 although the correlation between treated/untreated trials for height was much stronger
240 than for yield. HFN and grain protein content were positively correlated to each other,
241 but negatively correlated to treated yield, treated plant height and days to ripening.

242

243 **Examining Trait Trends**

244 We next examined the temporal pattern across the 15 years of trials to highlight linear
245 trends in traits due to breeding progress. For this, we regressed the EVM for each
246 variety on its year of first entry to the NVPT which is directly related to its year of
247 release. This regression likely captures temporal pattern of breeding progress as
248 successive releases of varieties are expected to outperform previous releases in one or
249 more traits. We observed linear increase for yield between 2002 - 2017 in both the
250 treated and untreated trials (Figure 3A - B). The rate of yield increase in the untreated
251 trial was significantly higher than in the treated trials (rate difference = 0.093
252 tonnes/ha/year, $P < 0.0001$). Conversely, grain protein content and HFN showed small

but significantly decrease over time ($P < 0.001$ and 0.03, respectively; Figure 3C - D). We also observed a significant delay in days to ripening over the same period ($P = 0.004$, Figure 3E). Changes in plant height (treated and untreated) and specific weight were not significant ($P = 0.31 – 0.51$, Figure 3F - H) suggesting stable trends.

UK wheat varieties are classified into four main end-use groups as described by the UK Flour Millers (www.ukflourmillers.org). These include the UK Flour Group 1 – 4, hereafter referred to as UFG1-4. The UFG1 and UFG2 varieties have superior grain quality (grain protein content and HFN) and are used for breadmaking. UFG3 varieties are often used for biscuits and cakes, whereas UFG4 varieties usually have high yield potential but inferior grain quality and are mainly used for animal feed. As yield and protein content are important measures for these end-use classifications, we examined how the temporal trends observed for these traits varied for the different end-use groups. Expectedly, UFG4 varieties showed higher yield while the bread making varieties (UFG1-2) show higher grain protein content (Figure 4A - B). All end-use groups showed a significant increasing yield trend across time and the rates of increase were not significantly different between the end-use groups ($P = 0.263 - 0.885$; Figure 4C). UFG2 and UFG4 varieties showed a significant and comparable decline in grain protein content over time (Figure 4D) while changes in protein content of UFG1 and UFG3 varieties were non-significant (Figure 4D).

Yield, protein content, specific weight, but not HFN, are stable in UK environments

Using a modified Finlay Wilkinson (FW) regression (Lian and de los Campos 2016) for measuring genotype x environment interaction, we examined the stability of yield and end-use quality traits across the trial years (Figure 5, Table S3). Only 95 varieties that were trialed in three or more years were included in this analysis. FW regression measures the stability of variety performance across different environments by regressing individual variety trait means on the environmental effect (Finlay and Wilkinson 1963). FW regression coefficient close to 1 suggests average varietal stability in which variety performance is consistent with environment effect i.e. variety performs poorly in bad environments and well in good environments. Larger values suggest below average stability i.e. higher environmental sensitivity.

Yield was stable across years in most UK wheat varieties (regression coefficients close to 1, Figure 5A). Similarly, most of the varieties examined showed high stability in protein content and specific weight, with bread-making varieties stably producing grains with above median protein and specific weights (Figure 5B). HFN, on the other hand, showed varying FW coefficients ranging from -0.28 (KWS Barrel) to 6.03 (Hyperion). More than 83% of the 95 UK wheat varieties examined have FW coefficient > 2 for HFN suggesting below-average stability. Figure 5B shows the HFN performance of three varieties with different FW coefficients: KWS_Barrel, Hyperion, and Napier with (FW coefficient of 1.02). Napier consistently showed low HFN values in all the years it was trialed. On the other hand, Hyperion with the highest FW co-efficient, showed extreme HFN phenotypes - very low HFN value in Low-HFN years and very high HFN value in

high-HFN years suggesting high environmental sensitivity. KWS_Barrel's HFN performance was fairly constant irrespective of the environments it was trialed.

Post-2002 UK wheat varieties belong into four distinct population groups

Using the Axiom35k SNP array (Allen *et al.* 2017) we genotyped 139 varieties including a subset of those trialed between 2002 - 2017 (104) and additional historic UK wheat cultivars. After quality filtering (described in Materials and Methods), we selected 4298 high quality markers dataset (Table S4) including 1715, 1781 and 778 markers on the A, B and D sub-genomes, respectively (Table S5). Using these genotypic data, we examined the population structure within the UK wheat collection. DAPC analysis revealed four distinct population groups (Pop1-4; Figure 6A, Table S6). Using Helium for pedigree visualization (Shaw *et al.* 2014), we could trace the modern founder parents for three (Pop1, 2 and 4) of the four population groups. Pop1 contains 19 varieties, of which 15 (79%) have Cadenza in their pedigree, consistent with Cadenza being an important parent for Pop1. Pop2 comprises 27 varieties, 20 (74%) of which contain Claire in their pedigree. Pop4 includes 30 varieties, 28 (93%) of which trace their pedigree to Robigus suggesting Robigus as an important parent for this group (Figure 6B). Pop3 is the largest group with 63 varieties with a more diverse pedigree structure. Using a subset of 111 varieties with both genotype and end-use group information (Figure S2), we examined the association between the population groups

and end-use groups (Figure S3). The “Claire” (Pop2) and “Robigus” (Pop4) population groups only contain UFG3 and UFG4 varieties used for biscuit/cakes and feeds, respectively. While the “Cadenza” (Pop1) population group mostly (71%) contain UFG1 and UFG2 varieties used for breadmaking.

Using NVPT Data for Trait Mapping

We next examined the suitability of using the EVM obtained from the NVPT for trait mapping through a genome-wide association study (GWAS). To ascertain that our genotypic data and population composition are suitable for GWAS, we included data for the presence/absence of *Sm1* - a major locus known to underlie resistance to Orange wheat blossom midge (OWBM) in UK wheat varieties. As expected, we identified a major peak associated with OWBM resistance on wheat chromosome 2B (Figure S4A and B). This peak co-localizes with the physical position for *Sm1* (Walkowiak *et al.* 2020), supporting our *Sm1* marker information. Importantly, our GWAS analysis identified a region on the short arm of chromosome 6A with significant marker trait association (MTA) for days to ripening (Figure 7A - B). The days to ripening MTA region contain two markers, AX-94549511 and AX-94710688, located in an interval (73.5 – 86.5 Mbp) containing the *NAM-A1* gene (TraesCS6A02G108300; 77.1 Mbp) that is associated with variation in senescence in European wheat cultivars (Cormier *et al.* 2015). Days to ripening was significantly different between the allele groups of marker AX-94710688 which has the highest significance score (Figure 7C).

DISCUSSIONS

Yield is an important driver of linear trends

Using historical data from UK NVPT we examined phenotypic trends in winter wheat varieties trialed between 2002 – 2017. Our analysis highlights a linear increase for yield (treated and untreated) and days to ripening, and a linear decrease in protein content and HFN. Given that the model used to analyze this data adjusted for variation arising from locations across years, and that agronomic practices are largely consistent in the NVPT, this linear trend can be attributed mostly to genetic improvement of varieties over time. Mackay *et al.* (2011) similarly attributed 88% of yield increase in cereals crops in the UK from 1982 – 2007 to genetic improvement. Yield is the most important determinant of grain market value; as such the linear increase in yield is consistent with concerted breeding efforts to improve yield under UK wheat growing conditions. In addition to the overall yield trend, we also observed consistent and similar linear increases in yield in all the four UK Flour Groups (UFG1 – 4). This further highlight yield as the main breeding target for varietal development (and adoption into the RL) irrespective of their target end-use groups.

We observed that the rate of yield increase in untreated trials (152 kg/ha/year) is significantly ($p < 0.0001$) higher than in treated trials (60 kg/ha/year) across the 15-year period. Mackay *et al.* (2011) similarly observed the same pattern between 1982 – 2007 and argued that this pattern is due to loss of disease resistance by some varieties during the trial period examined. Varieties progressively lose resistance over time

(Meikle and Scarisbrick 1994) and consequently variety performance declines with time. This mean that under untreated trial conditions, newly introduced varieties with 'intact' disease resistance will outperform a portion of previously released varieties whose disease resistance have 'broken down'. This differential loss of disease resistance will further increase the variation in variety yield performance in untreated trials in addition to the variation arising from non-disease related genetic factors observed in treated trials. In other words, there is an "upward bias" in variety effects for the yield observed in untreated trials as described by Mackay *et al.* (2011).

Based on the rationale described above, it would be expected that a sudden loss of resistance in a large proportion of varieties due to the emergence of a more virulent pathogen race would result in a marked upward bias in variety effect estimates. This is what we observed when we compared yield trends before and after the emergence of the yellow rust (*Puccinia striiformis*) "Warrior" race in 2011 (Hubbard *et al.* 2015). The rate of yield increase in untreated trials significantly ($P < 0.001$) increased three-fold from 123 kg/ha/year before the emergence of the "Warrior" race to 372 kg/ha/year after the emergence of the "Warrior" race (Figure 8). During the same time, the rate of yield increase was significantly ($P = 0.2697$) comparable in the treated trial before and after the emergence of the "Warrior" race (Figure 8). The use of historical data in this study allowed us to identify this trend and thus highlight the importance of such datasets for dissecting the effect of important events in a national cropping history such as change in disease epidemics.

It is also interesting to speculate that the higher rate of yield increase observed in the untreated trials indirectly suggests that newer varieties contain new sources of genetic

385 resistance that improve their performance over older varieties at a rate greater than
386 observed in the treated trials. This is likely not accidental, but points to concerted efforts
387 by breeders to introduce more effective source of genetic resistance into UK wheat. The
388 improved genetic resistance profile of newer varieties narrows the yield gap observed
389 between the treated and untreated trials. We cannot, however, rule out the fact that this
390 narrower yield gap might be due to less disease pressure in recent years. A more
391 detailed genetic characterization will be needed to accurately describe the genetic
392 resistance profile of UK wheat varieties.

393 Concomitant with the yield increase, there has been a decrease in grain protein content
394 from 2002 - 2017 which reflects the well-established antagonistic relationship between
395 yield and protein content (Figure S5; Simmonds 1995). Unlike for yield, linear trends
396 were not consistent across the four end-use groups. While we identified an overall
397 significant decrease in grain protein content over time, this was not observed in the
398 UFG1 varieties that are used for breadmaking (Figure 4). UFG2 varieties which also
399 have breadmaking potential, however, showed significant decrease over time just like
400 the UFG4 varieties used for animal feed. The decline in UFG2 varieties grain protein
401 content may be due to the fact that this group comprise varieties that did not
402 consistently meet the higher grain quality (in particular protein content) requirement for
403 UFG1 and were downgraded to UFG2 . The fact that our analysis captures expected
404 trait (yield, protein content and HFN) differences in end-use groups (Figure 4A - B,
405 Figure S6A - B) suggests that the linear mixed effect model adopted is appropriate to
406 handle the incomplete design of the NVPT and to examine phenotype trends within
407 each end-use group.

The multidimensional (year and location) nature of the NVPT also allows for examining varietal adaptability across multiple environments. We observed year-to-year stability in yield and protein content in most of the varieties irrespective of their end-use group. This is likely attributable to the fact that we mainly examined data from RL trials that are comprised of varieties which had been previously screened for distinctness, uniformity, and stability during National Listing trials. Despite this 'pre-screening', almost all the varieties show high environmental sensitivity for HFN (FW coefficient: -0.28 to 6.03). Sjoberg et al (2020) similarly obtained a wide range of FW coefficient for HFN in 133 varieties trialed across three years in the Pacific Northwest of the US.

HFN is inversely related to α-amylase activity within the grain. High α-amylase activity caused by incidences of pre-harvest sprouting (PHS) and/or pre-maturity amylase (PMA) reduce the bread-making potential of wheat grains. Both PHS and PMA are known to be highly environmental dependent: PHS is induced by wet raining conditions during harvest maturity while PMA is mostly caused by low or high temperature shock around grain physiological maturity (Joe *et al.* 2005; Mares and Mrva 2014). The environmental conditions required to induce PHS and PMA occur infrequently from year to year making it difficult for breeders to screen for these traits under field conditions. In addition, both traits are controlled by many genes most of which have small effects making marker assisted selection (MAS) for HFN stability difficult. Within the last decade, progress has been made in identifying genes with major effects on PHS including *TaMFT* and *TaMKK3*-A (Nakamura *et al.* 2011; Torada *et al.* 2016). We also previously showed the effect of *TaMMK3-A* in reducing PHS in UK germplasm

(Shorinola *et al.* 2016) and developed markers to facilitates its use in breeding (Shorinola *et al.* 2017). The availability of markers for major genes controlling PHS now makes it possible to apply MAS for improving HFN. However, selection for PMA resistance remains a major challenge because the conditions that induces PMA varies between varieties (Liu *et al.* 2021)

Population structure within UK winter wheat germplasm

Our analysis reveals that three modern wheat varieties largely contribute to the development of winter wheat varieties released in the UK between 2002 - 2017. These include Cadenza (Pop1), Claire (Pop2) and Robigus (Pop4), which were themselves released in 1992, 1999, 2005, respectively. Together, 51% of the 114 varieties that were first trialed between 2002 – 2017 were derived from either Cadenza, Claire, and/or Robigus. Based on pedigree visualization, Robigus (and Pop4 varieties) appears to be a more recent introduction to the UK (Figure S7) suggesting that new gene pools are being introduced into the UK wheat breeding landscape. Since its introduction Robigus has made significant contribution to UK wheat pedigree. Fradgley et al (2019) identified Robigus as the second most used parents in UK breeding, next to Capelle Desprez. We also observed a clear association between the population groups and end use groups. Pop2 and Pop4 varieties, mostly derived from Claire and Robigus which are themselves UFG3 varieties, both contain only UFG3 (biscuit) and UFG4 (feed) varieties. Pop1 varieties, which are mostly derived from Cadenza - a UFG2 variety, mostly contain UFG1 and UFG2 (breadmaking) varieties. One probable explanation for this association is that breeders tend to make crosses with varieties from the same end-use groups to

ensure that the gene combinations underlying the traits in the target end-use groups are preserved in their progenies (Simon Berry 2021, personal communication). This suggests that the choice of parents is an important determinant of the end-use class of varieties.

Due to the type (gene-based SNP) and limited number of markers used, we acknowledge the limitation of this study to more precisely define the population groups represented in UK winter bread wheat collection to a high resolution. Brinton et al. (2020) demonstrated the inadequacy of array-based genotyping chips to precisely define haplotype groups due to their gene-centric design. Scaffold-level assemblies are now available for important UK wheat varieties including representatives of Pop1, Pop3 and Pop4 (Cadenza, Claire and Robigus; Walkowiak *et al.* 2020). These genome assemblies can be combined with high-density genotyping or re-sequencing data to more precisely define the populations groups of wheat varieties grown in the UK.

Historical data could be valuable for trait mapping

We identified significant marker-trait association (MTA) peaks spanning a gene (*NAM-A1*) that have been previously associated with natural variation in a trait of agronomic interest.. Cormier et al. (2015) identified a C/T missense SNP in the NAC domain and A/- frame-shift deletion in *NAM-A1* leading to a truncated protein from a worldwide wheat collection and suggested functional roles for these polymorphisms. Harrington et al., (2019) showed that missense mutations in the NAC domain of *NAM-A1* result in delayed peduncle and flag leaf senescence. Similarly, Avni et al (2014) showed that loss of function *NAM-A1* mutants showed significant delay in senescence. Given the

large interval covered by the MTA peaks for days to ripening on chromosome 6A (73.4 Mbp – 86.5 Mbp, ~140 genes) we cannot rule out the possibility that other gene(s) underly this days to ripening effect. Nonetheless, the co-localization of our GWAS peak with a known locus for the target trait highlights the usefulness of this historical dataset for quantitative trait mapping.

Beside the MTA for days to ripening, we did not identify strong MTA for the other traits. This might be due to the fact that many of the major genes controlling these traits have been mostly fixed in the UK wheat population examined, and that the population size used in our study is not large enough to pick up minor effect and/or minor allele frequency gene(s). Also, while the phenotyping conditions used in the NPVT might be representative of UK farming conditions, they might not always be best suited for trait mapping. An example is the application of plant growth regulators in the trials to prevent lodging (by reducing plant height) but this might mask the effect of height genes. Despite these limitations, our work demonstrates that national trials data can be valuable for examining trait trends, stability, and genetic architecture.

Acknowledgments

We thank the Agriculture and Horticulture Development Board (AHDB) for making the trial data publicly available. We also thank Dr Chinyere Ekine for useful discussion on statistical analyses and Dr Simon Berry for suggestions to the manuscript.

Funding information

This research is supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) Designing Future Wheat program (BB/P016855/1) and a Royal Society FLAIR award (FLR\R1\1918500) to OS.

Conflicts of Interest

The authors declare no conflict of interest.

REFERENCES

Adamski, N. M., P. Borrill, J. Brinton, S. A. Harrington, C. Marchal *et al.*, 2020 A roadmap for gene functional characterisation in crops with large genomes: Lessons from polyploid wheat (C. S. Hardtke, Ed.). eLife 9: e55646.

Allen, A. M., M. O. Winfield, A. J. Burridge, R. C. Downie, H. R. Benbow *et al.*, 2017 Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (Triticum aestivum). Plant Biotechnology Journal 15: 390–401.

Avni, R., R. Zhao, S. Pearce, Y. Jun, C. Uauy *et al.*, 2014 Functional characterization of GPC-1 genes in hexaploid wheat. Planta 239: 313–324.

518 Ayalew, H., P. W. Tsang, C. Chu, J. Wang, S. Liu *et al.*, 2019 Comparison of TaqMan,
519 KASP and rhAmp SNP genotyping platforms in hexaploid wheat. PLOS ONE 14:
520 e0217222.
521 Bolser, D., D. M. Staines, E. Pritchard, and P. Kersey, 2016 Ensembl Plants: Integrating
522 Tools for Visualizing, Mining, and Analyzing Plant Genomics Data, pp. 115–140 in *Plant*
523 *Bioinformatics: Methods and Protocols*, edited by D. Edwards. Methods in Molecular
524 Biology, Springer, New York, NY.
525 Brinton, J., R. H. Ramirez-Gonzalez, J. Simmonds, L. Wingen, S. Orford *et al.*, 2020 A
526 haplotype-led approach to increase the precision of wheat breeding. Communications
527 Biology 3: 712.
528 Cheng, H., J. Liu, J. Wen, X. Nie, L. Xu *et al.*, 2019 Frequent intra- and inter-species
529 introgression shapes the landscape of genetic variation in bread wheat. Genome
530 Biology 20: 136.
531 Cormier, F., M. Throude, C. Ravel, J. L. Gouis, M. Leveugle *et al.*, 2015 Detection of
532 NAM-A1 Natural Variants in Bread Wheat Reveals Differences in Haplotype Distribution
533 between a Worldwide Core Collection and European Elite Germplasm. Agronomy 5:
534 143–151.
535 Crossa, J., J. Burgueño, S. Dreisigacker, M. Vargas, S. A. Herrera-Foessel *et al.*, 2007
536 Association Analysis of Historical Bread Wheat Germplasm Using Additive Genetic
537 Covariance of Relatives and Population Structure. Genetics 177: 1889–1913.
538 Finlay, K., and G. Wilkinson, 1963 The analysis of adaptation in a plant-breeding
539 programme. Aust. J. Agric. Res. 14: 742–754.

Fradgley, N., K. A. Gardner, J. Cockram, J. Elderfield, J. M. Hickey *et al.*, 2019 A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. PLOS Biology 17: e3000071.

Gardiner, L.-J., L. U. Wingen, P. Bailey, R. Joynson, T. Brabbs *et al.*, 2019 Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency. Genome Biology 20: 69.

Harrington, S. A., L. E. Overend, N. Cobo, P. Borrill, and C. Uauy, 2019 Conserved residues in the wheat (Triticum aestivum) NAM-A1 NAC domain are required for protein binding and when mutated lead to delayed peduncle and flag leaf senescence. BMC Plant Biology 19: 407.

He, F., R. Pasam, F. Shi, S. Kant, G. Keeble-Gagnere *et al.*, 2019 Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. Nature Genetics 51: 896–904.

Howe, K. L., B. Contreras-Moreira, N. De Silva, G. Maslen, W. Akanni *et al.*, 2020 Ensembl Genomes 2020—enabling non-vertebrate genomic research. Nucleic Acids Research 48: D689–D695.

Hubbard, A., C. M. Lewis, K. Yoshida, R. H. Ramirez-Gonzalez, C. de Vallavieille-Pope *et al.*, 2015 Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. Genome Biology 16: 23.

IWGSC, R. Appels, K. Eversole, N. Stein, C. Feuillet *et al.*, 2018 Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361: eaar7191.

562 Joe, A. F. T. W., R. W. Summers, G. D. Lunn, M. D. Atkinson, and P. S. Kettlewell,
563 2005 Pre-maturity α-amylase and incipient sprouting in UK winter wheat, with special
564 reference to the variety Rialto. Euphytica 143: 265–269.
565 Jombart, T., and I. Ahmed, 2011 adegenet 1.3-1: new tools for the analysis of genome-
566 wide SNP data. Bioinformatics 27: 3070–3071.
567 Jordan, K. W., S. Wang, Y. Lun, L.-J. Gardiner, R. MacLachlan *et al.*, 2015 A haplotype
568 map of allohexaploid wheat reveals distinct patterns of selection on homoeologous
569 genomes. Genome Biology 16: 48.
570 Krasileva, K. V., H. A. Vasquez-Gross, T. Howell, P. Bailey, F. Paraiso *et al.*, 2017
571 Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci USA 114: E913.
572 Lenth, R. V., 2016 Least-Squares Means: The R Package lsmeans. Journal of
573 Statistical Software 69: 33.
574 Lian, L., and G. de los Campos, 2016 FW: An R Package for Finlay–Wilkinson
575 Regression that Incorporates Genomic/Pedigree Information and Covariance Structures
576 Between Environments. G3: Genes|Genomes|Genetics 6: 589.
577 Liu, C., K. M. Tuttle, K. A. Garland Campbell, M. O. Pumphrey, and C. M. Steber, 2021
578 Investigating conditions that induce late maturity alpha-amylase (LMA) using
579 Northwestern US spring wheat (Triticum aestivum L.). Seed Science Research 1–9.
580 Mackay, I., A. Horwell, J. Garner, J. White, J. McKee *et al.*, 2011 Reanalyses of the
581 historical series of UK variety trials to quantify the contributions of genetic and
582 environmental factors to trends and variability in yield over time. Theor Appl Genet 122:
583 225–238.

584 Mares, D. J., and K. Mrva, 2014 Wheat grain preharvest sprouting and late maturity
585 alpha-amylase. Planta 240: 1167–1178.
586 Meikle, S. M., and D. H. Scarisbrick, 1994 Cereal Breeding and Varietal Testing. British
587 Food Journal 96: 11–16.
588 Nakamura, S., F. Abe, H. Kawahigashi, K. Nakazono, A. Tagiri *et al.*, 2011 A Wheat
589 Homolog of MOTHER OF FT AND TFL1 Acts in the Regulation of Germination. Plant
590 Cell 23: 3215–3229.
591 Pozniak, C. J., J. M. Clarke, and F. R. Clarke, 2012 Potential for detection of marker–
592 trait associations in durum wheat using unbalanced, historical phenotypic datasets. Mol
593 Breeding 30: 1537–1550.
594 Rasheed, A., and X. Xia, 2019 From markers to genome-based breeding in wheat.
595 Theoretical and Applied Genetics 132: 767–784.
596 Rosyara, U. R., W. S. De Jong, D. S. Douches, and J. B. Endelman, 2016 Software for
597 Genome-Wide Association Studies in Autopolyploids and Its Application to Potato. Plant
598 Genome 9:.
599 Scott, M. F., N. Fradgley, A. R. Bentley, T. Brabbs, F. Corke *et al.*, 2020 Limited
600 haplotype diversity underlies polygenic trait architecture across 70 years of wheat
601 breeding. bioRxiv 2020.09.15.296533.
602 Semagn, K., R. Babu, S. Hearne, and M. Olsen, 2014 Single nucleotide polymorphism
603 genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology
604 and its application in crop improvement. Molecular Breeding 33: 1–14.
605 Semenov, M. A., 2009 Impacts of climate change on wheat in England and Wales.
606 Journal of The Royal Society Interface 6: 343–350.

607 Shaw, P. D., M. Graham, J. Kennedy, I. Milne, and D. F. Marshall, 2014 Helium:
608 visualization of large scale plant pedigrees. BMC Bioinformatics 15: 259.
609 Shorinola, O., B. Balcárková, J. Hyles, J. F. G. Tibbits, M. J. Hayden *et al.*, 2017
610 Haplotype Analysis of the Pre-harvest Sprouting Resistance Locus Phs-A1 Reveals a
611 Causal Role of TaMKK3-A in Global Germplasm. Front Plant Sci 8: 1555–1555.
612 Shorinola, O., N. Bird, J. Simmonds, S. Berry, T. Henriksson *et al.*, 2016 The wheat
613 Phs-A1 pre-harvest sprouting resistance locus delays the rate of seed dormancy loss
614 and maps 0.3 cM distal to the PM19 genes in UK germplasm. Journal of Experimental
615 Botany 67: 4169–4178.
616 Silvey, V., 1981 The contribution of new wheat, barley and oat varieties to increasing
617 yield in England and Wales 1947-78.
618 Simmonds, N. W., 1995 The relation between yield and protein in cereal grain. Journal
619 of the Science of Food and Agriculture 67: 309–315.
620 Sjoberg, S. M., A. H. Carter, C. M. Steber, and K. A. Garland-Campbell, 2020
621 Unraveling complex traits in wheat: Approaches for analyzing genotype × environment
622 interactions in a multienvironment study of falling numbers. Crop Science 60: 3013–
623 3026.
624 Su, G., P. Madsen, M. S. Lund, D. Sorensen, I. R. Korsgaard *et al.*, 2006 Bayesian
625 analysis of the linear reaction norm model with unknown covariates. J Anim Sci 84:
626 1651–1657.
627 Sweeney, D. W., J. Sun, E. Taagen, and M. E. Sorrells, 2019 Genomic Selection in
628 Wheat, pp. 273–302 in *Applications of Genetic and Genomic Research in Cereals*,
629 edited by T. Miedaner and V. Korzun. Woodhead Publishing.

630 Torada, A., M. Koike, T. Ogawa, Y. Takenouchi, K. Tadamura *et al.*, 2016 A Causal
631 Gene for Seed Dormancy on Wheat Chromosome 4A Encodes a MAP Kinase Kinase.
632 Curr Biol 26: 782–787.
633 Trnka, M., S. Feng, M. A. Semenov, J. E. Olesen, K. C. Kersebaum *et al.*, 2019
634 Mitigation efforts will not fully alleviate the increase in water scarcity occurrence
635 probability in wheat-producing areas. Sci Adv 5: eaau2406.
636 Walkowiak, S., L. Gao, C. Monat, G. Haberer, M. T. Kassa *et al.*, 2020 Multiple wheat
637 genomes reveal global variation in modern breeding. Nature 588: 277–283.
638 Wang, S., D. Wong, K. Forrest, A. Allen, S. Chao *et al.*, 2014 Characterization of
639 polyploid wheat genomic diversity using a high-density 90 000 single nucleotide
640 polymorphism array. Plant Biotechnology Journal 12: 787–796.
641 Wilkinson, P. A., A. M. Allen, S. Tyrrell, L. U. Wingen, X. Bian *et al.*, 2020 CerealsDB—
642 new tools for the analysis of the wheat genome: update 2020. Database 2020:.
643 Winfield, M. O., A. M. Allen, A. J. Burridge, G. L. A. Barker, H. R. Benbow *et al.*, 2016
644 High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary
645 gene pool. Plant Biotechnology Journal 14: 1195–1206.
646 Yang, W., H. Feng, X. Zhang, J. Zhang, J. H. Doonan *et al.*, 2020 Crop Phenomics and
647 High-Throughput Phenotyping: Past Decades, Current Challenges, and Future
648 Perspectives. Molecular Plant 13: 187–214.
649
650
651
652

FIGURE LEGEND

Figure 1: Distribution of 162 NVPT locations used in this study (2002 and 2017). The number of field sites within each county and unitary authority are indicated in colour.

Figure 2: Phenotype correlation between yield, adaptation and grain quality traits. EVM were derived for each variety from the NVPT conducted between 2002 and 2017. Only significant correlations ($P < 0.05$) are indicated. Positive and negative correlations are indicated with the blue and red circles, respectively, with the size and colour intensity of the circles representing the magnitude of the correlation.

Figure 3: Temporal Trait Trend in UK Winter Wheat. Scatter plot showing changes in yield in the treated (A) and untreated trial (B), protein content (C), HFN (D), days to ripening (E), specific weight (F), plant height in treated (G) and untreated (H) trials. Blue dots represent individual varieties. For each trait, the EVM for each variety is regressed against the first year of entry in the 2002 -2017 trials. The solid line shows the regression line of the linear model and is coloured red if significant ($P < 0.05$). The shaded region defines the confidence interval. The regression equation is shown within each plot. The EVM data used for these plots are in Table S2.

Figure 4: Temporal Trait Trend by End-use Groups. (A-B) Violin plots showing distribution for yield (A) and protein content (B) for the different end-use groups. The solid lines represent the mean of the distribution and the black letters show Tukey statistical comparison between the groups. Groups that are statistically similar share the same letter. (C-D) Scatter plot showing changes in yield (C) and protein content (D) for each end-use group of UK winter wheat. Each dot represents a variety while the colors

675 of the dots represent the end use groups (UK Flour Group 1-4). For each trait, the EVM
676 for each variety is regressed against the first year of entry in the 2002 -2017 trials. The
677 solid lines are the regression line of the linear model. The regression line equation for
678 each group is shown. UFG1, UFG2, UFG3 and UFG4 are represented by the red,
679 green, gray and peach dots, lines and text, respectively.

680 **Figure 5**: Phenotype Stability by End-use Group. Scatter plot showing stability for
681 treated yield, protein content, specific weight and HFN for UK winter wheat varieties
682 across the 15 years of trials (2002 – 2017). The y-axis represents the Finlay Wilkinson
683 (FW) coefficient which specifies expected change in performance per unit change in
684 environment (year) effect. Varieties with above median performance are in the shaded
685 region. The solid line indicates stable performance in all environments i.e. b + 1 = 1
686 (Lian and De los Campos, 2016). Datapoints for three varieties whose HFN
687 performance are further illustrated in Figure 5B are labeled. (B) Plot of HFN
688 performance of varieties with lowest, highest and stable (~1) FW coefficient against the
689 estimated environment year effect. The dashed lines present a constant slope of 1.

690 **Figure 6**: Population structure of UK winter wheat varieties using DAPC analysis. (A)
691 The representative variety for each population group (Pop) is indicated except for Pop2
692 which consists of a more diverse pedigree. (B) Pedigree structure for Pop4 “Robigus”.
693 The number in the inset represent varieties: (1) Qplus (2) Torch (3) Viscount (4)
694 Conqueror (5) Leeds (6) Lear (7) Zulu (8) Gravitas (9) Twister (10) Britannia (11) Invicta
695 (12) Warrior (13) Cougar (14) KWS Croft (15) KWS Target (16) Oakley (17) Jorvik (18)
696 Panacea (19) Tuxedo (20) Icon (21) Horatio (22) KWS Gator (23) KWS Santiago (24)
697 RGT Scrummage (25) Reflection (26) Energise (27) KWS Kerrin. The population groups

698 are represented by teal (Pop1), yellow (Pop2), purple (Pop3), and red (Pop4) circles,
699 whereas gray circles represent varieties which were not genotyped in this study.

700 **Figure 7**: (A) Manhattan plot for days to ripening using EVM derived from the 2002 –
701 2017 NVPT of UK winter wheat varieties. The Bonferroni threshold is indicated with a
702 dotted line. The seven wheat chromosome groups are indicated on the X-axis and each
703 homoeologous sub-genome is coloured in red (A genome), gray (B) or yellow (D). (B)
704 QQplot showing expected and observed distribution of –log (p values). (C) Allele effect
705 of the marker showing the highest significant marker trait association for days to
706 ripening.

707 **Figure 8**: Yield comparison between treated and untreated trials before and after the
708 emergence of the “warrior” yellow rust race. Scatter plot showing changes in yield in
709 treated (light blue) and untreated trials (dark blue) before (unshaded region) and after
710 (shaded region) the emergence of the “warrior” yellow rust race. The EVM for each
711 period are regressed separately against the first year of entry into the NVPT trials for
712 each variety. The solid lines are the regression lines. The regression equations are
713 shown at the bottom corners of the plot.

714

715

716

717

718

719

720

721

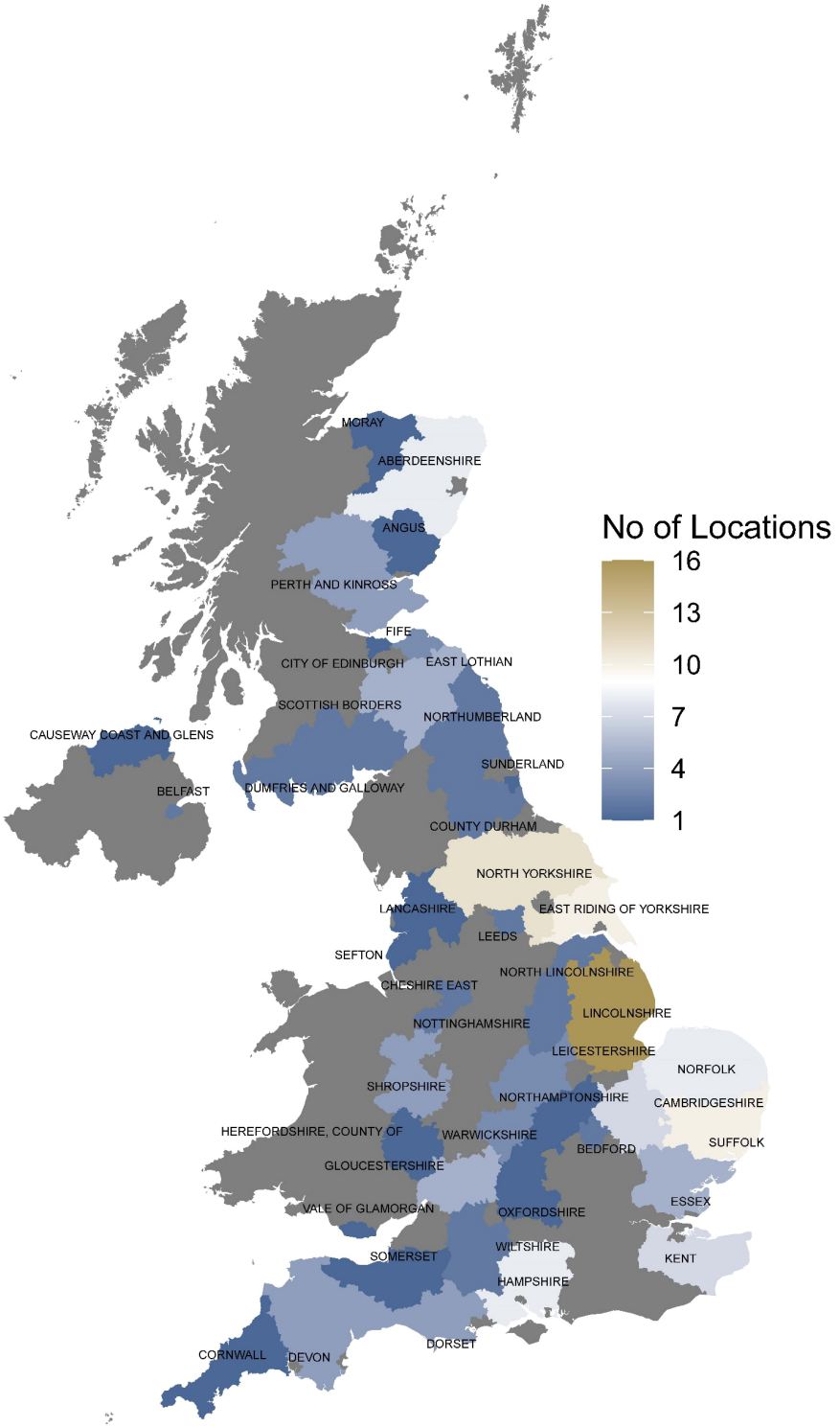
722

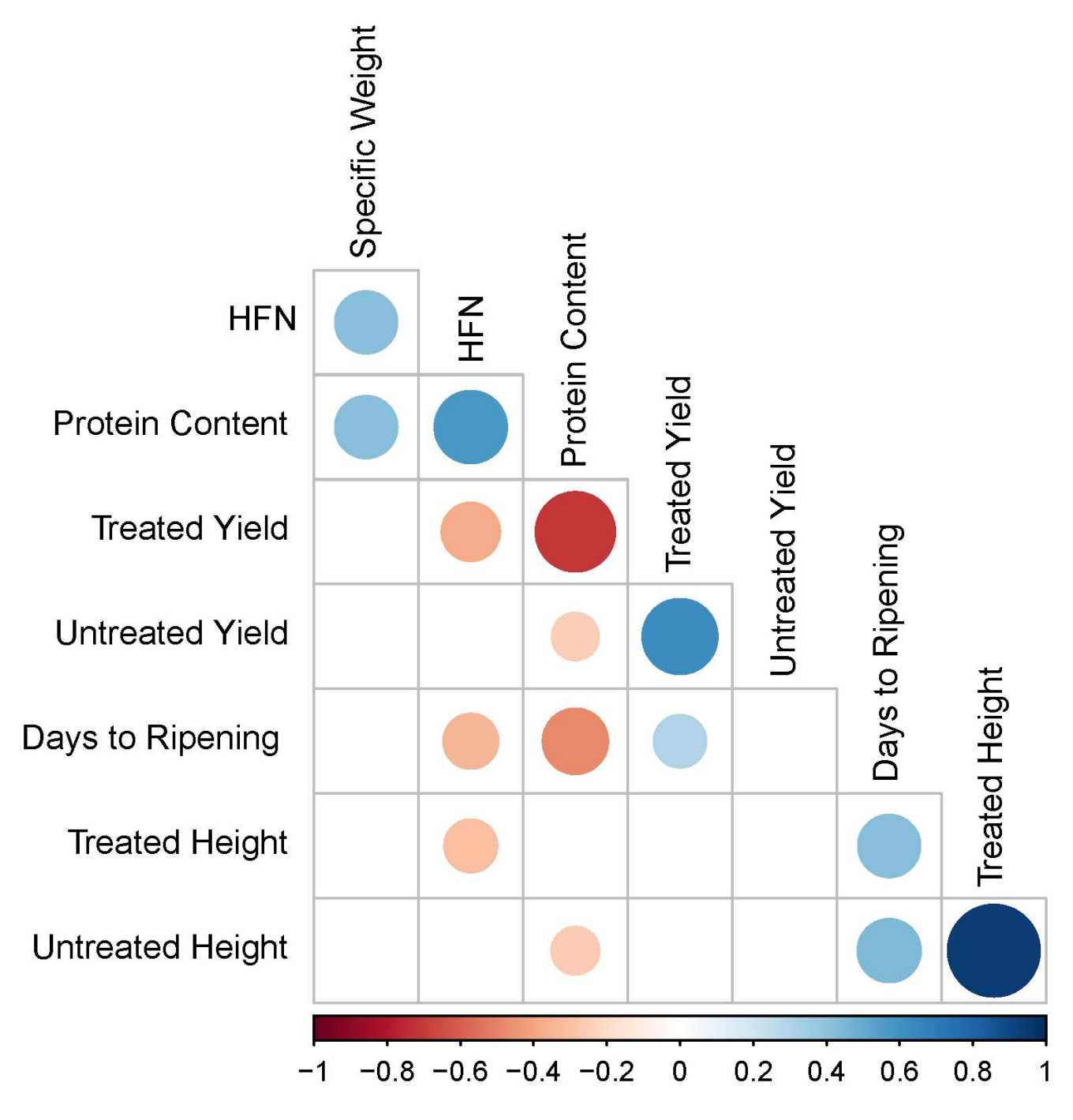
723 TABLE 1: NUMBER OF VARIETIES, SITES, AND YEARS OF TRIALS FOR THE UK NVPT BETWEEN

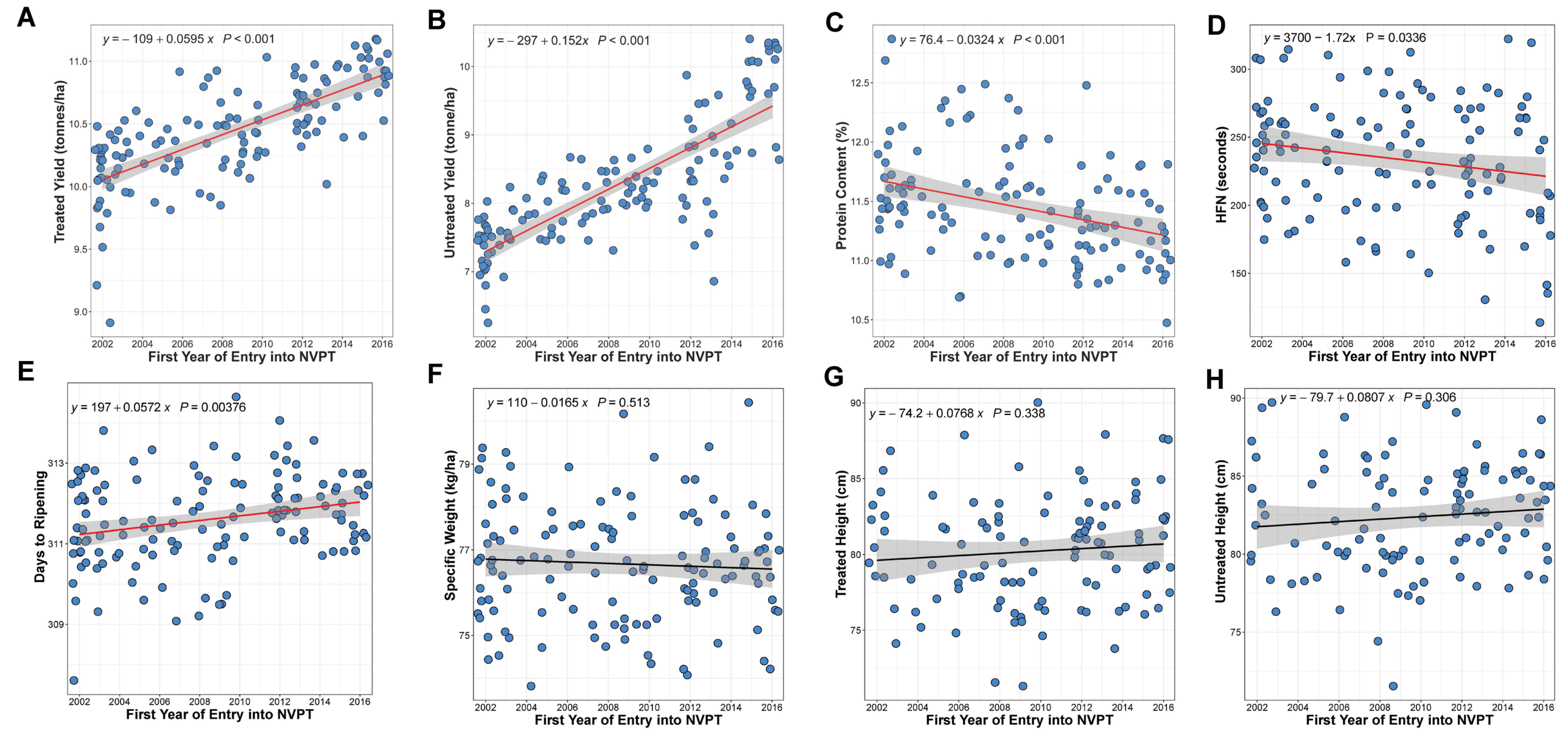
724 2002-2017

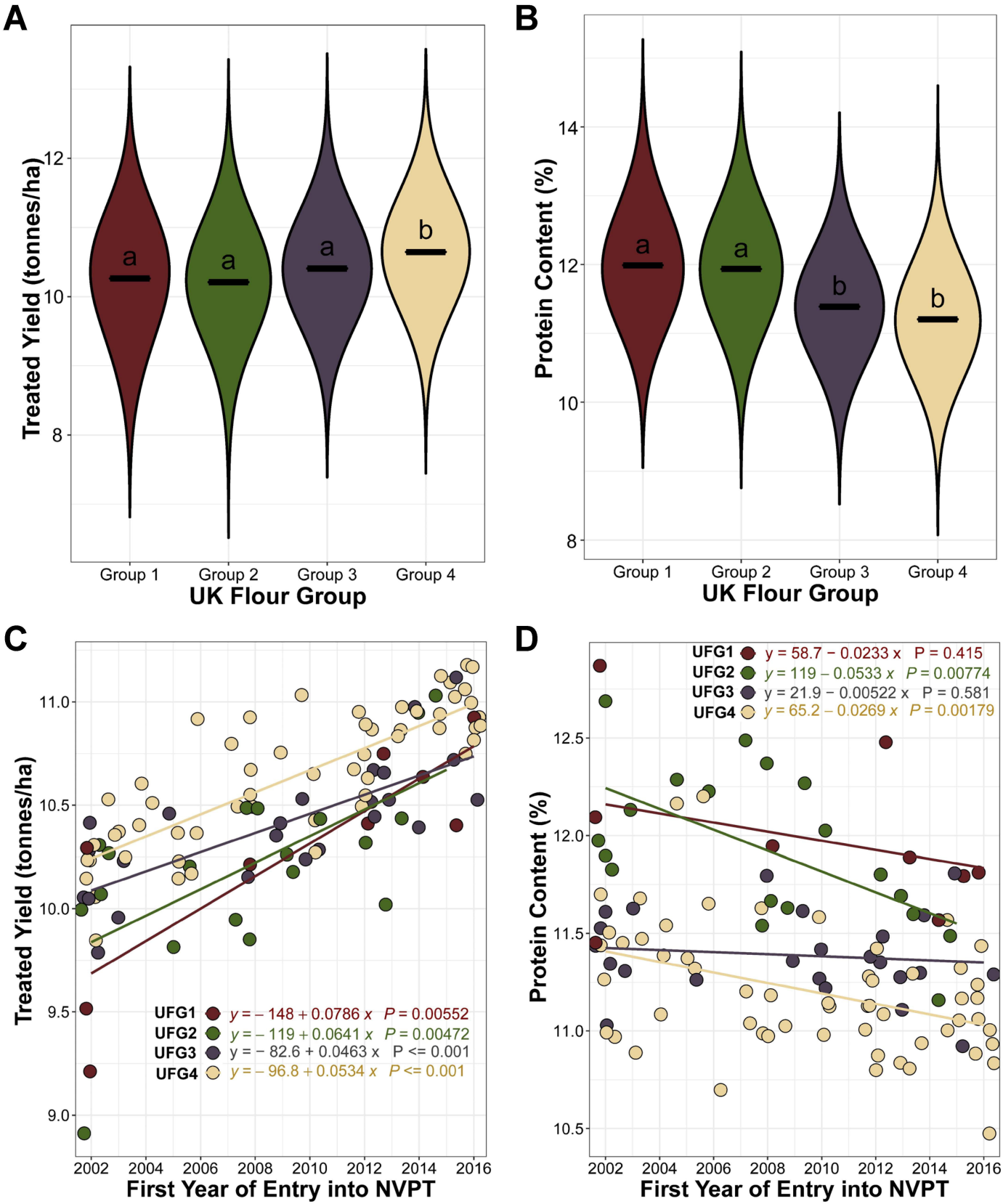
Trait	Varieties*	Trial Locations++	Trial Years	Year x Location Combinations	Total Observations
Treated yield	133	158	15	410	13080
Untreated yield	131	53	15	124	4156
Protein content	128	99	15	230	7142
Days to ripening	133	108	15	247	7977
HFN	128	99	15	227	7154
Specific weight	129	99	15	231	7091
Treated height	108	74	11	171	3905
Untreated height	107	30	11	75	1647

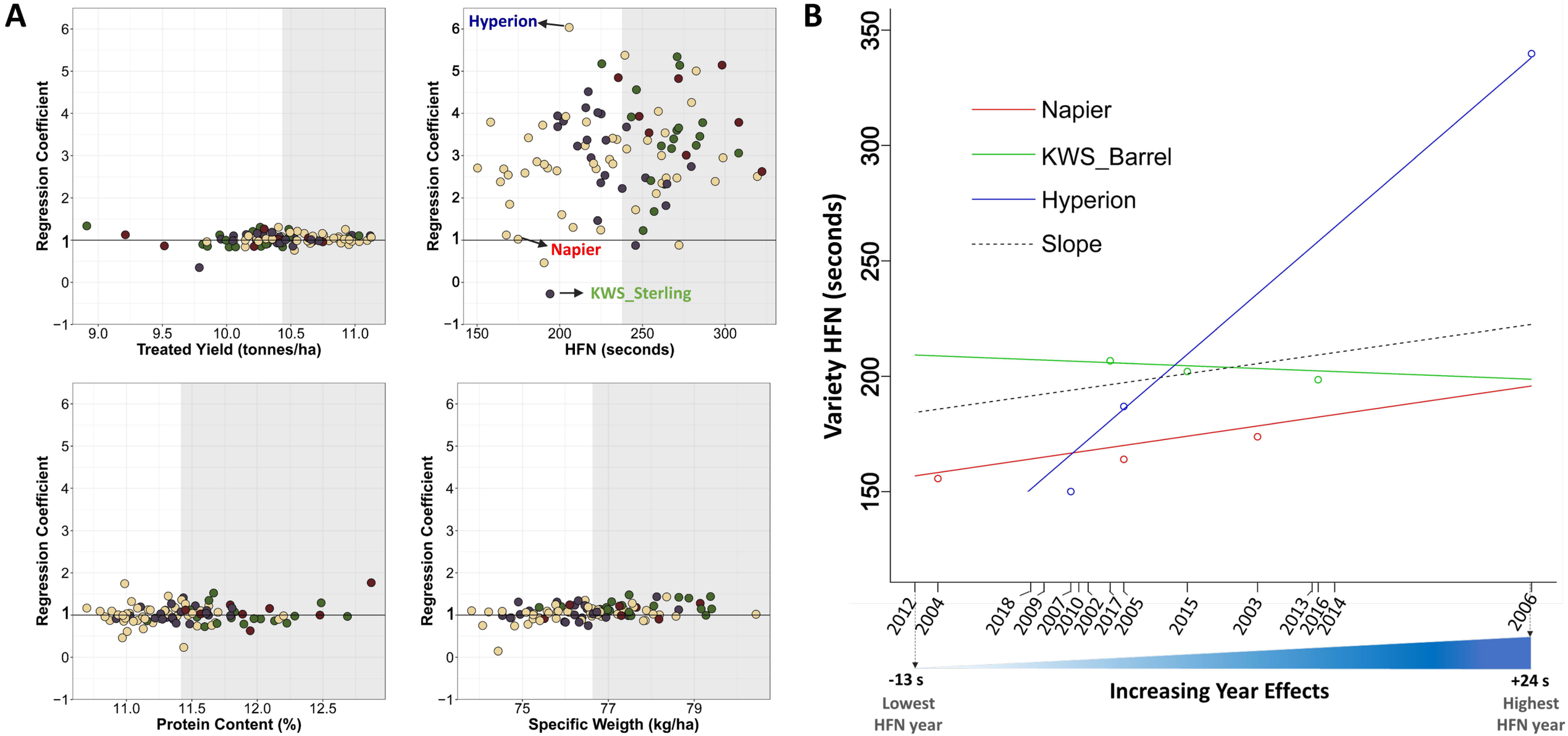
725 *Not all varieties were tested for each trait, and in each year and location.
726 ++Some locations were used in more than one year.
727

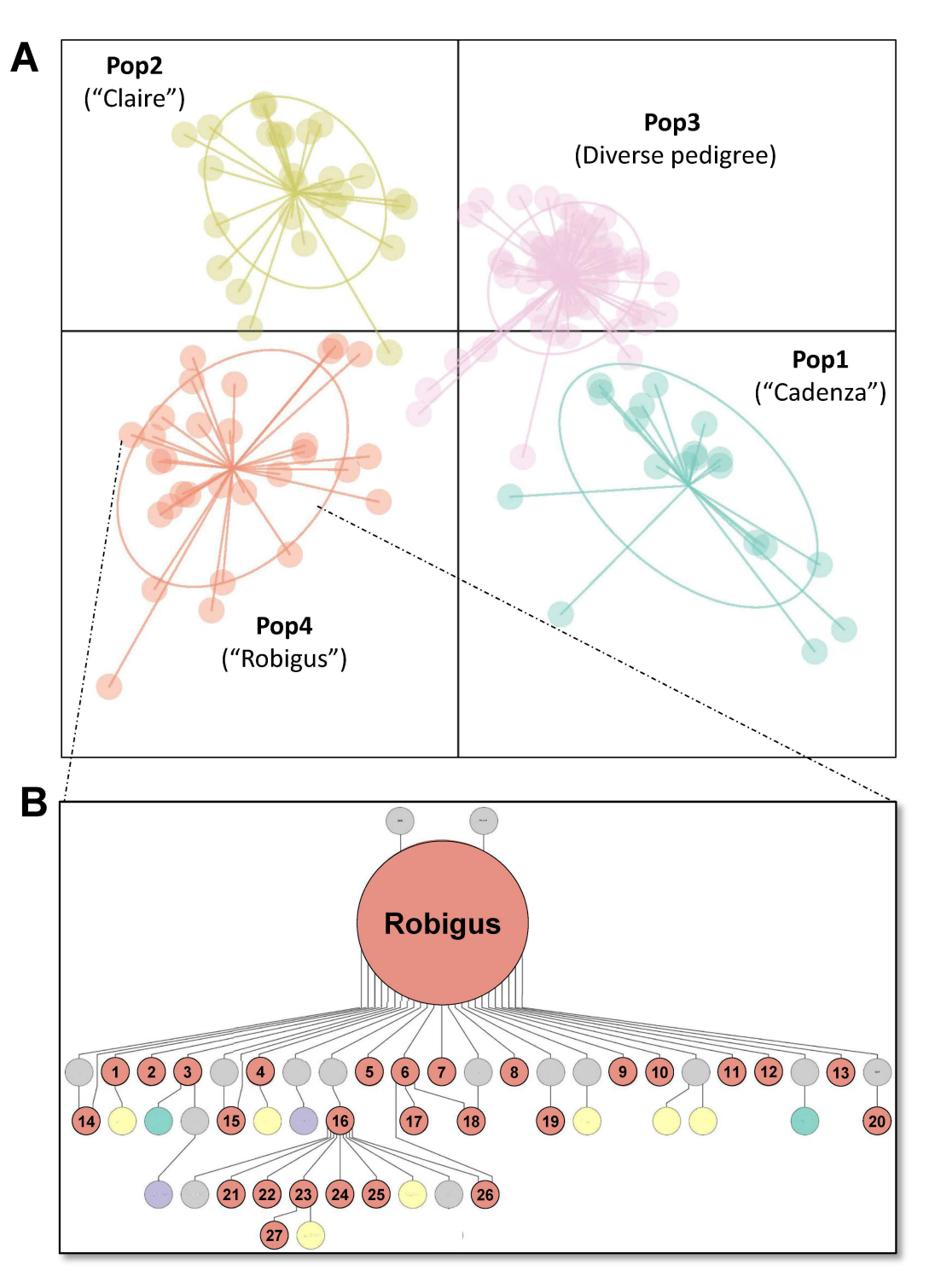

No of Locations
16
13
10
7
4
1
MORAY
ABERDEENSHIRE
ANGUS
PERTH AND KINROSS
FIFE
CITY OF EDINBURGH
EAST LOTHIAN
SCOTTISH BORDERS
NORTHUMBERLAND
CAUSEWAY COAST AND GLENS
SUNDERLAND
BELFAST
DUMFRIES AND GALLOWAY
COUNTY DURHAM
NORTH YORKSHIRE
LANCASHIRE
EAST RIDING OF YORKSHIRE
LEEDS
SEFTON
NORTH LINCOLNSHIRE
CHESHIRE EAST
LINCOLNSHIRE
NOTTINGHAMSHIRE
LEICESTERSHIRE
NORFOLK
SHROPSHIRE
NORTHAMPTONSHIRE
CAMBRIDGESHIRE
HEREFORDSHIRE, COUNTY OF
WARWICKSHIRE
BEDFORD
SUFFOLK
GLOUCESTERSHIRE
ESSEX
VALE OF GLAMORGAN
OXFORDSHIRE
WILTSHIRE
KENT
SOMERSET
HAMPSHIRE
DORSET
CORNWALL
DEVON

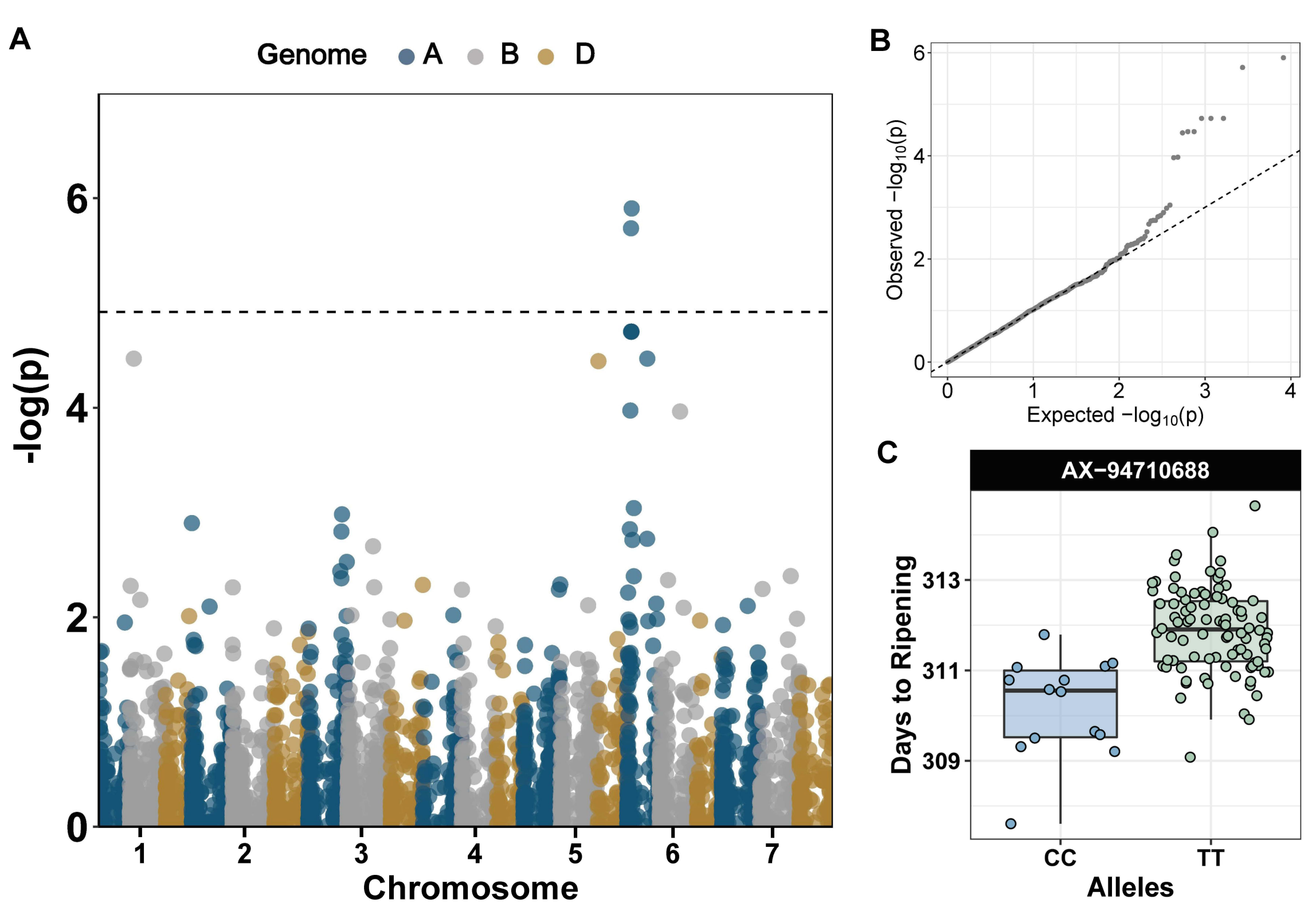

Specific Weight
HFN
Protein Content
Treated Yield
Untreated Yield
Days to Ripening
Treated Height
Untreated Height
HFN
Protein Content
Treated Yield
Untreated Yield
Days to Ripening
Treated Height
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1

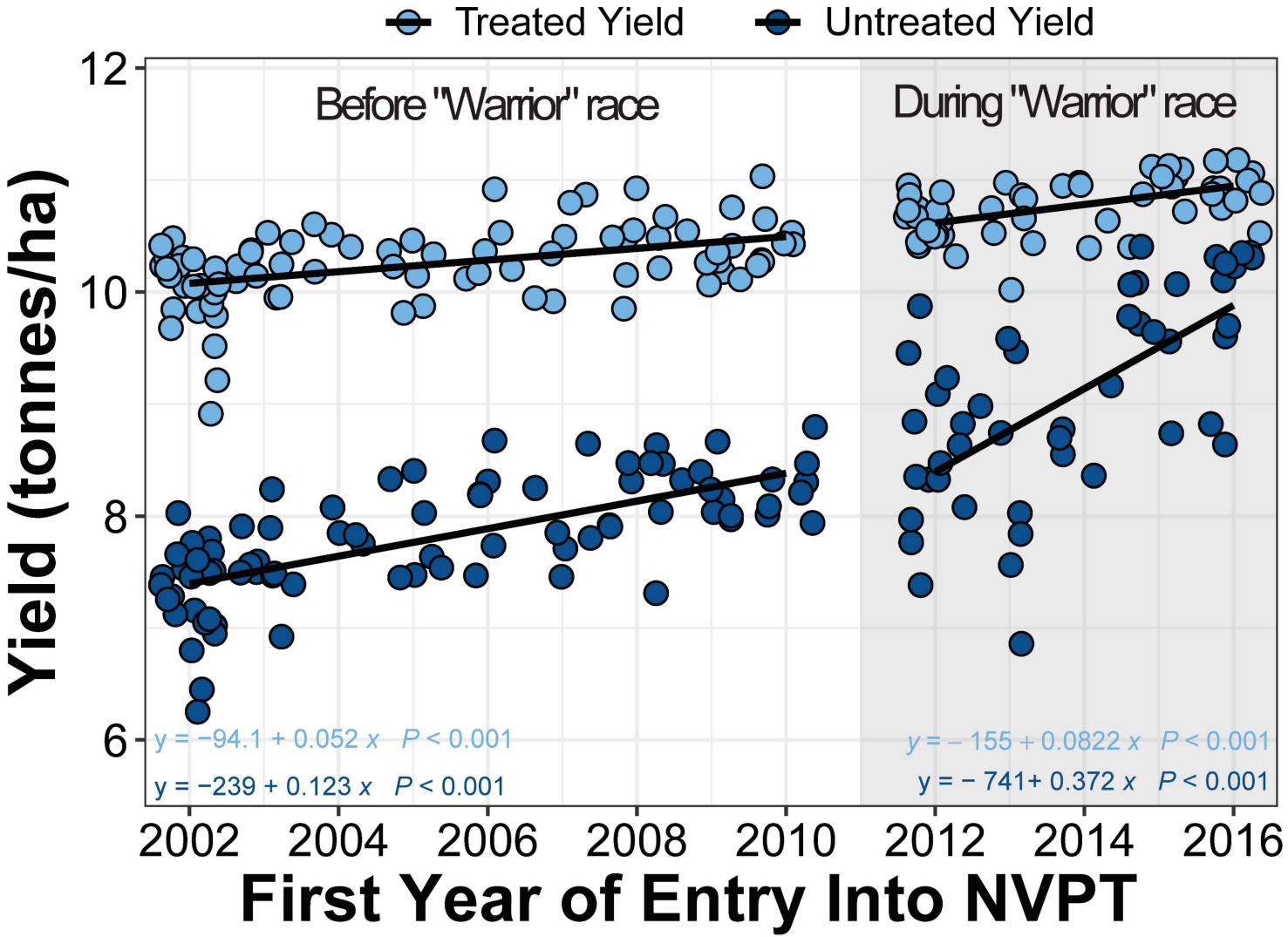

A
y = – 109 + 0.0595 x P < 0.001
Treated Yield (tonnes/ha)
First Year of Entry into NVPT
B
y = – 297 + 0.152x P < 0.001
Untreated Yield (tonne/ha)
First Year of Entry into NVPT
C
y = 76.4 – 0.0324 x P < 0.001
Protein Content (%)
First Year of Entry into NVPT
D
y = 3700 – 1.72x P = 0.0336
HFN (seconds)
First Year of Entry into NVPT
E
y = 197 + 0.0572 x P = 0.00376
Days to Ripening
First Year of Entry into NVPT
F
y = 110 – 0.0165 x P = 0.513
Specific Weight (kg/ha)
First Year of Entry into NVPT
G
y = – 74.2 + 0.0768 x P = 0.338
Treated Height (cm)
First Year of Entry into NVPT
H
y = – 79.7 + 0.0807 x P = 0.306
Untreated Height (cm)
First Year of Entry into NVPT


A
Treated Yield (tonnes/ha)
12
10
8
a
a
a
b
Group 1
Group 2
Group 3
Group 4
UK Flour Group
B
Protein Content (%)
14
12
10
8
a
a
b
b
Group 1
Group 2
Group 3
Group 4
UK Flour Group
C
Treated Yield (tonnes/ha)
11.0
10.5
10.0
9.5
9.0
UFG1 y = – 148 + 0.0786 x P = 0.00552
UFG2 y = – 119 + 0.0641 x P = 0.00472
UFG3 y = – 82.6 + 0.0463 x P <= 0.001
UFG4 y = – 96.8 + 0.0534 x P <= 0.001
2002
2004
2006
2008
2010
2012
2014
2016
First Year of Entry into NVPT
D
Protein Content (%)
12.5
12.0
11.5
11.0
10.5
UFG1 y = 58.7 – 0.0233 x P = 0.415
UFG2 y = 119 – 0.0533 x P = 0.00774
UFG3 y = 21.9 – 0.00522 x P = 0.581
UFG4 y = 65.2 – 0.0269 x P = 0.00179
2002
2004
2006
2008
2010
2012
2014
2016
First Year of Entry into NVPT


A
Regression Coefficient
Treated Yield (tonnes/ha)
HFN (seconds)
Hyperion
Napier
KWS_Sterling
Protein Content (%)
Specific Weigth (kg/ha)
B
Variety HFN (seconds)
Napier
KWS_Barrel
Hyperion
Slope
2012
2004
2018
2009
2007
2010
2002
2017
2005
2015
2003
2013
2016
2014
2006
-13 s
Lowest
HFN year
Increasing Year Effects
+24 s
Highest
HFN year


A
Pop2
("Claire")
Pop3
(Diverse pedigree)
Pop1
("Cadenza")
Pop4
("Robigus")
B
Robigus
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27


A
Genome A B D
-log(p)
0
2
4
6
1
2
3
4
5
6
7
Chromosome
B
Observed −log10(p)
Expected −log10(p)
0
2
4
6
0
1
2
3
4
C
AX−94710688
Days to Ripening
309
311
313
CC
TT
Alleles


Treated Yield
Untreated Yield
Before "Warrior" race
During "Warrior" race
Yield (tonnes/ha)
12
10
8
6
y = −94.1 + 0.052 x P < 0.001
y = −239 + 0.123 x P < 0.001
y = – 155 + 0.0822 x P < 0.001
y = − 741+ 0.372 x P < 0.001
2002
2004
2006
2008
2010
2012
2014
2016
First Year of Entry Into NVPT