1 Historical changes in the contents and compositions of fibre components and polar
2 metabolites in white wheat flour.

3 Alison Lovegrove[1*], Till K Pellny[1], Kirsty L Hassall[1], Amy Plummer[1], Abigail Wood[1], Alice
4 Bellisai[2], Alexandra Przewieslik-Allen[3], Amanda J. Burridge[3], Jane L. Ward[2] and Peter R
5 Shewry[1]
6
7 [1]Plant Science Department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK.
8 [2]Computational and Analytical Sciences Department, Rothamsted Research, Harpenden,
9 Hertfordshire AL5 2JQ, UK
10 [3] Life Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ
11 * Corresponding author. Tel: (44) 1582 763133. FAX: (44) 1582 763010. E-mail:
12 alison.lovegrove@rothamsted.ac.uk
13

14 Abstract

15 Thirty-nine UK adapted wheat cultivars dating from between 1790 and 2012 were grown in
16 replicated randomised field trials for three years, milled, and white flour analysed for the
17 contents of dietary fibre components (arabinoxylan and β-glucan) and polar metabolites
18 (sugars, amino acids, organic acids, choline and betaine) to determine whether the
19 composition had changed due to the effects of intensive breeding. The concentrations of
20 components varied between study years, indicating strong effects of environment.
21 Nevertheless, some trends were observed, with the concentrations of arabinoxylan fibre and
22 soluble sugars (notably sucrose, maltose and fructose) increasing and most amino acids
23 (including asparagine which is the precursor of acrylamide formed during processing)
24 decreasing between the older and newer types. The concentration of betaine, which is
25 beneficial for cardio-vascular health, also increased. The study therefore provided no
26 evidence for adverse effects of intensive breeding on the contents of beneficial components
27 in wheat flour.

28 Introduction

29 Scientific plant breeding has been immensely successful in increasing the yield and
30 improving the performance of wheat. For example, Mackay et al[1] calculated that about 88%
31 of the gain in yield of winter wheat in the UK between 1981 and 2007, from about 6 to 8
32 tonnes, is attributable to genetic improvement. However, modern cultivars have lower
33 genetic diversity than older cultivars and land races, particularly in the A and B
34 subgenomes[2]. It has also been suggested that modern plant breeding, with emphasis on
35 high yield (which effectively reflects starch content) and, in the case of wheat, on gluten

36 protein content, may have impacts on grain composition which result in negative effects on
37 health[3,4]. In fact, Kasarda[5] reported a decline in the protein content of wheats grown in the
38 Northern Plains of the USA, and we have reported similar deceases in the protein contents
39 of UK wheats[6].
40 Because starch constitutes about 80% of the grain, a decrease in protein content with
41 increasing yield would be expected and is often ascribed to "yield dilution". Yield dilution may
42 also contribute to the decreases in the mineral micronutrients (iron and zinc) that have
43 occurred since the introduction of dwarfing genes in the 1960s[7,8]. However, other effects of
44 dwarfing genes, for example on mineral nutrient uptake and partitioning, may also have
45 contributed[6].

46 Effects of intensive breeding on other bioactive components are less clear. Shewry et al[9]
47 compared the contents of phytochemicals in a global collection of 146 bread wheat
48 genotypes in relation to their dates of registration. No clear relationships were identified, but
49 only a single set of samples were analysed and these were grown on the same site in
50 Hungary, which was outside the area of adaptation of many of the lines. However,
51 comparisons of smaller numbers of "old and recent" adapted cultivars showed no difference
52 in the total contents of phenolics in durum or bread wheats, although the composition was
53 more diverse in the older cultivars[10,11]. A study of eight modern and 7 older Italian durum
54 wheats cultivars showed no differences in contents of arabinoxylan and β-glucan in
55 wholemeal and semolina, but higher arabinoxylan solubility in modern cultivars[12]. These
56 studies have largely focused on wholemeal samples, which are richer in bioactive
57 components. However, the most widely consumed foods in many countries, including the
58 UK, are produced from white flour and hence the relevance to human health of analyses
59 carried out on whole grains is debatable. We have therefore determined historical trends in
60 the composition of white flour of bread wheat, by comparing 39 varieties which are adapted
61 to the UK where they have been grown commercially over the past 200 years.

62 **Results**

63 **Selection and genetic diversity of wheat cultivars.** A series of 39 wheat cultivars
64 (including cultivars and earlier land races, all referred to as cultivars here) was selected to
65 represent the diversity in wheat grown in the UK since 1790 (Figure 1). All were either winter
66 type, or winter-hardy, and hence routinely grown as winter wheats. They include 9 out of the
67 11 cultivars grown in the Broadbalk continuous winter wheat experiment (which has been
68 grown at Rothamsted since 1843[13]). All selected cultivars had been grown commercially in
69 the UK, with some being regarded as "landmark varieties", and all except 4 were bred in the
70 UK (Figure 1). Hence, they can be regarded as well adapted to the UK climate. The cultivars

71 are divided into three groups, which are colour coded and represent stages in the
72 development of wheat breeding. Group 1 comprises 9 cultivars which were released
73 between 1790 and 1916. During this period selections were made from landraces and from
74 populations from early crosses, but selection was empirical without an understanding of
75 genetic mechanisms. The second group comprises 13 cultivars released between 1935 and
76 1972, which represent the increasing application of scientific theory to wheat breeding. The
77 third group comprises 17 cultivars released between 1980 and 2012 which represent the
78 products of modern breeding technologies (with investment in wheat breeding being
79 stimulated in the UK by an increased demand for homegrown wheat following the accession
80 into the European Union in 1973). The major scientific advance during this period was the
81 introduction in the 1970s of the "green revolution" dwarfing genes which increase the harvest
82 index and hence yield. Consequently, the *Rht2* dwarfing gene is present in 13 of these
83 cultivars and the *Rht1* dwarfing gene in one (Figure 1). In addition, increasing use was made
84 of "alien introgression", for example, Xi19, Cadenza, Robigus and Crusoe all contain
85 introgressions from *Triticum dicoccoides*, and of new technologies to increase the efficiency
86 of breeding (such as doubled haploid production). These three groups are therefore termed
87 "empirical selection", "early breeding" and "modern breeding". The pedigrees of the cultivars,
88 where known, are shown in Supplementary Figure S1.

89 The broad genetic relationships between the cultivars were initially determined using the
90 Axiom Wheat HD Genotyping Array, comprising 819,571 SNP markers, and the data
91 analysed by principal component analysis (PCA) (Figure 2). Comparison of PCs 1 and 2,
92 which accounted for 8.66% and 6.58% of the total variation, respectively, showed that the
93 cultivars released since 1980 were more closely related to each other than those released
94 before 1980. Two cultivars are clearly separated from the others: April Bearded (1838) and
95 Apollo (1984) (labelled in Figure 2). In both cases the separations may result from
96 introgressions. Wider studies have shown that April Bearded has DNA in common with
97 *Triticum aestivum* ssp. *compactum* (also called club wheat) while Apollo has DNA in
98 common with rye (presumably derived from Triticale which is present in the pedigree, see
99 Supplementary Figure S1) (authors' unpublished results). This analysis, and plots of further
100 PCs (PC3 5.85%, PC4 5.1%, PC5 3.88%, PC6 3.68%) (not shown), indicate that the recent
101 cultivars are less genetically diverse that the older cultivars.

102 **Crop and grain phenotyping.** Decreases in the height of wheat cultivars grown in the UK
103 over the past century are well-documented[14] and have been reported previously for the
104 cultivars in the present study[6]. The use of small experimental plots, and the difficulty in
105 harvesting some of the older cultivars, precluded the determination of grain yield, or of grain
106 number which is one of the two determinants of yield. However, grain weight and kernel

107 diameter both showed significant differences between cultivar groups ($F_{2,225}$=106.41, $p < 0.001$ and $F_{2,225}$= 115.18, $p < 0.001$). Furthermore, although these group differences differed between experimental years (significant interaction effects, $F_{4,225}$= 35.93, $p < 0.001$ and $F_{4,225}$= 16.32, $p < 0.001$), both grain weight and kernel diameter tended to be higher in the more recent (post-1935) cultivars (Table 1, Supplementary Table S1, Supplementary Figure S2). Similar studies have shown that the grain weight of wheat cultivars did not increase in the USA over the period 1919-1987[15] or in Canada between 1947 and 1992[16]. Grain hardness also showed significant differences between groups ($F_{2,225}$= 504.36, $p < 0.001$) with more recent groups, on average, having a higher grain hardness (Table 1, Supplementary Table S1). This is probably due to increased emphasis on breeding for breadmaking quality.

118 **Flour composition.** In order to determine whether there were differences in the composition of white flour, which accounts for about 90% of the flour used for breadmaking in the UK (http://www.nabim.org.uk/statistics/), white flour fractions were prepared and analysed for two groups of components: dietary fibre (arabinoxylan and β-glucan) and polar metabolites (comprising mainly amino acids, sugars and small oligosaccharides, arabinogalactan peptide (AGP), choline and betaine). These components were selected because wheat is the major source of dietary fibre in the UK diet[17] and the metabolites include components which contribute to health. The full datasets for individual components are given in Supplementary Table S2.

127 Statistical analysis of the data using ANOVA shows highly significant differences between individual cultivars for all measured variables (Supplementary Table S3). Furthermore, strong differences are observed between the three groups which, with 2 exceptions (choline and arabinose equivalents in AGP), are greater than the differences observed between the individual cultivars within each group. Although it is clear that environmental factors affect the content and composition of fibre and metabolites, this interaction is small compared with the main effect of genotype.

134 In order to focus on differences between the cultivars, mean values for eight individual components or groups of components over the three years are summarised in Table 2; Figure 3 and Supplementary Table S3.

137 Table 2 (rows 1 and 2) and Figure 3 (parts A and B) show the amounts of AX and β-glucan, the two major dietary fibre components in white flour. Both components show significant differences between cultivar groups ('Group' column), ($F_{2,221}$= 212.17, $p < 0.001$ and $F_{2,221}$= 36.47, $p < 0.001$). Moreover, the difference between groups is consistent over study years for AX, (row 1, 'Time.Group' column) ($F_{4,221}$= 1.06, $p = 0.3764$), despite changes in the

142 differences within each group over the study years ('Time.Group.Cultivar' column) ($F_{72,221}$= 2.34, p <0.001). By contrast, the group differences for β-glucan are not consistent over study
144 years (row 2, 'Time.Group' column) ($F_{4,221}$= 5.71, p < 0.001). The associated cultivar means
145 (shown in Table 2, Figure 3 parts A and B and Supplementary Table S4) highlight the
146 stronger trend in the amount of AX (clearly higher in recent cultivars) compared to the
147 amount of β-glucan, which varies more and shows weaker trends except for being low in the
148 early cultivars.

149 Changes in the structure and composition of the dietary fibre fraction were also studied by
150 Principal Component Analysis (PCA), comparing the proportions of arabinoxylan
151 oligosaccharides (AXOS) released by digestion of AX with endoxylanase and of gluco-
152 oligosaccharides (G3 and G4 GOS) released by digestion of β-glucan with lichenase (β-
153 glucanase). Figure 4 (parts A and B) compares PCs 1 and 2, which together account for
154 59% of the total variation. Although the three groups of cultivars clearly overlap, partial
155 separation is observed with the modern cultivars clustered in the left-hand part of the
156 separation. The loadings plot (Figure 4B) shows that the separation along PC1 is associated
157 with Xyl5 and Xyl3 (positively associated with modern breeding) and XA2+3XX,
158 XA3A2+3XX, XA3A3XX, G3 and G4 (positively associated with empirical selection and early
159 breeding).

160 Table 2 (rows 3 and 4) and Figure 3 (parts C and D) show the concentrations of total free
161 amino acids and free asparagine, the latter being of interest to grain processors as it is a
162 precursor, and usually the limiting factor, for acrylamide formation during wheat
163 processing[18]. The concentrations of both show downward trends which are associated with
164 strong statistical differences between groups (('Group' column, $F_{2,221}$= 44.03, p < 0.001
165 and$F_{2,221}$= 27.02, p < 0.001)). These are consistent over study years for total free amino
166 acids ('Time.Group' column, $F_{4,221}$= 1.65, p = 0.1630) but less so for free asparagine
167 ('Time.Group' column, $F_{4,221}$= 3.34, p = 0.0111). Similar trends are observed for the
168 concentrations of most of the other individual amino acids (Supplementary Tables S3 and
169 S4).

170 By contrast, the concentrations of total monosaccharides and small oligosaccharides (called
171 total carbohydrates in Figure 3) are generally higher in recent cultivars (Table 2 row 5 and
172 Figure 3 part E) with strong statistical differences being observed between groups (although
173 these are not consistent over study years) (('Group' column, $F_{2,221}$= 241.95, p < 0.001),
174 ('Time.Group' column, $F_{4,221}$= 4.30, p = 0.0023)). This fraction comprises sucrose, raffinose,
175 maltose, glucose, fructose, galactose and arabinose, with sucrose, maltose and fructose
176 showing the clearest increases with time (Supplementary Tables S3 and S4).

Table 1 (row 6) and Figure 3 (part F) show total methyl donors, which comprises choline and betaine (glycine betaine), and shows highly significant differences between cultivar groups (‘Group’ column), ($F_{2,221}$= 152.41, $p < 0.001$). Betaine is generally present at 10-fold higher concentrations in wheat than choline[19], and the increase in total methyl donors shown in Table 2 and Figure 3 is due to higher contents of betaine in more recent cultivars, with little change in the contents of choline (Supplementary Tables S3 and S4).

Finally, Table 1 (rows 7 and 8) and Figure 3 (parts G and H) show total organic acids and the wheat arabinogalactan peptide (AGP). The organic acids comprise fumaric acid, succinic acid and malic acid. The concentrations of these components show strong differences between the groups of cultivars (‘Group’ column, $F_{2,22,1}$= 16.34, $p < 0.001$), albeit much weaker than for the previously discussed components. No clear trends (Figure 3, part G and Table 1) can be identified between the groups. AGP is a short (15 amino acid) peptide which is *o*-glycosylated, probably on three hydroxyproline residues[20]. It accounts for about 0.4% of white flour[21] and is readily fermented by faecal bacteria *in vitro*[22], indicating that it may have prebiotic properties *in vivo.* The concentration shown in Figure 3 (part H) is the mean of galactose and arabinose equivalents determined by NMR spectroscopy (Supplementary Table S3). Although significant differences between cultivar groups are observed (‘Group’ column, $F_{2,221}$= 9.53, $p < 0.001$), no clear trends across the groups of cultivars are evident.

To confirm the overall trends discussed above, the full datasets for all metabolites over the three years were compared by PCA analysis. Figure 4 (parts C and D) compares PCs 1 and 2, which together account for 63% of the total variation. Although there is overlap between the groups of cultivars based on release date, some separation between the older and most recent groups is observed (Figure 4, part C). The loadings plot (Figure 4, part D) shows that this separation is related to lower and higher concentrations of amino acids and sugars, respectively, in the most recent cultivars. This separation is confirmed by the individual ANOVAs (Supplementary Table S3), where highly significant group effects are seen, particularly in sucrose, tryptophan, betaine, fumaric acid and raffinose.

Discussion

The cultivars compared here were selected because they have been widely grown in the UK. Hence, the differences observed should not be related to their degree of adaptation. Nevertheless, all components measured were highly affected by the environment, as shown by the comparison of samples from three harvests shown in Supplementary Table S4. Therefore, in order to identify broad trends, it was decided to calculate the means of the

211 contents determined for the three years for individual cultivars, and then the means of three
212 groups of cultivars selected to represent different stages of wheat breeding. When this was
213 done, clear trends were observed for some components, as summarised in Figure 3, and
214 Table 2

215 The components measured included several which are considered to contribute to effects on
216 the health of consumers. The most important of these for most consumers is dietary fibre, as
217 bread provides about 20% of the total daily intake in the UK, and white bread about half of
218 this[17]. Hence, the higher contents of arabinoxylan, the major dietary fibre component, in
219 modern cultivars are particularly noteworthy. By contrast, fermentable sugars may have
220 beneficial or adverse effects. Raffinose and fructose have been defined as FODMAPs
221 (fermentable oligo-, di- and monosaccharides and polyols), a group of compounds which
222 have been implicated in causing discomfort in patients with irritable bowel syndrome (IBS)[23].
223 By contrast, these sugars and AGP may also have beneficial prebiotic effects in healthy
224 individuals. The biochemical basis for the increased concentrations of sugars is not known,
225 but it could relate to the higher levels of starch synthesis and accumulation. Similarly, the
226 lower concentrations of amino acids in the recent cultivars could relate to their lower content
227 of protein, which decreased from about 16.9% to 12.5% in the sample sets from years 1 and
228 2 (determined as N x 6.25 and reported by Shewry et al[6]).

229 Finally, betaine and choline are biosynthetically related components which are considered to
230 be beneficial for cardio-vascular health, by acting as methyl donors in the homocysteine
231 cycle[24]. Wheat is a particularly rich source of these compounds, which together account for
232 about 1.5 to 3 mg/g dry wt in wholemeal[19]. The increased concentration of betaine in the
233 samples could therefore contribute to greater health benefits.

234 The conclusion from this study is, therefore, that there is no evidence that the health benefits
235 of white flour from wheat grown in the UK have declined significantly over the past 200
236 years. In fact, increasing trends in several components, notably the major form of dietary
237 fibre (arabinoxylan) are observed. This is despite great increases in the yields of wheat
238 grown over this period. However, there are strong environmental effects on grain
239 composition which must therefore be taken into account when comparing the compositions
240 of grain samples.

241 Methods

242 **Plant material.** 39 bread wheat cultivars were selected to represent diversity in UK adapted
243 commercial wheats released and grown between 1790 and 2012 (Figure 1). These were
244 grown at Rothamsted Research in three replicate $1m^2$ plots for three successive seasons:

245 2013-2014, 2014-2015 and 2015-2016. Nitrogen was applied as ammonium nitrate at
246 210kg/Ha (2013-2014) or 150 kg/Ha (2014-2015, 2015-2016) with other inputs being
247 according to standard agronomic practice. Plots were staked where necessary and heads
248 harvested and threshed by hand. Grain was conditioned to 14% water content and milled
249 using a Chopin CD1 mill to give white flour.

250 **Genotyping.** The Axiom Wheat HD Genotyping Array (Thermo Fisher Scientific, Inc.,
251 Waltham, MA) (comprising 816,571 SNP markers) was used to genotype the 39 samples
252 using the Affymetrix GeneTitan (Thermo Fisher Scientific, Inc.) system according to the
253 procedure described by Affymetrix (Life Technologies, 2017). Allele calling was performed
254 using the Affymetrix proprietary software package Axiom Analysis Suite, following the Axiom
255 Best Practices Genotyping Workflow. A distance matrix was generated from the genotype
256 scores using R package SNPRelate[25]. The proportion of variance for the first six eigenvalues
257 was as follows: 8.66, 6.58, 5.85, 5.10, 3.88, 3.68. The first two eigenvalues accounting for
258 over 15% of the variance were plotted as a PCA plot.

259 **Arabinoxylan and β-glucan.** Enzymatic fingerprinting of AX was as described previously[26].
260 White flour was digested using a mixture of endoxylanase and lichenase (β-glucanase) to
261 release arabinoxylan oligosaccharides (AXOS) and gluco-oligosaccharides (GOS)
262 comprising 3 and 4 residues (G3, G4), respectively. These were separated using a
263 Carbopac PA-1 (Dionex) column with dimensions 2 mm × 250 mm and the flow rate of 0.25
264 mL/min based upon the original method of Ordaz-Ortiz et al.[27]. At least two technical
265 replicates of each biological replicate were analysed. The areas under the AXOS peaks
266 were combined to determine TOT-AX and under the G3 and G4 GOS peaks to give total β-
267 glucan (expressed in arbitrary units).

268 **NMR spectroscopy.** ^{1}H-NMR sample preparation was carried out according to the
269 procedures described previously[28,29]. Flour samples (30 mg) were extracted 80:20
270 D_2O:CD_3OD containing 0.05% d_4– trimethylsilylpropionate (TSP) (1ml) as internal standard.
271 ^{1}H-NMR spectra were acquired under automation at 300 °K using an Avance Neo
272 Spectrometer (Bruker Biospin, Coventry, UK) operating at 600.0528 MHz, equipped with a
273 cryoplatform and a 5mm triple resonance inverse (TCI) probe. Spectra were collected using
274 a water suppression pulse sequence (zgpr) with a 90° pulse and a relaxation delay of 5 s.
275 Each spectrum was acquired using 16 scans of 65,536 data points with a spectral width of
276 7143 Hz. Spectra were automatically Fourier-transformed using an exponential window with
277 a line broadening value of 0.5 Hz. Phasing and baseline correction were carried out within
278 the instrument software. ^{1}H chemical shifts were referenced to d_4-TSP at δ0.00.

279 ^{1}H-NMR spectra were automatically reduced, using Amix (Analysis of MIXtures software, 280 BrukerBiospin), to ASCII files containing integrated regions or ‘buckets’ of equal width (0.01 281 ppm). Spectral intensities were scaled to the d_4-TSP region (δ0.05 to –0.05). The ASCII file 282 was imported into Microsoft Excel for the addition of sampling/treatment details. Signal 283 intensities for characteristic spectral regions for 29 major metabolites were extracted via 284 comparison to library spectra of known standards run in the same solvent system using 285 equivalent NMR data acquisition and processing parameters.

286

287 **Statistical methods**

288 Analysis of variance (ANOVA) was used to assess the effect of variety differences over the 3 289 experiments. Each field trial was an independent randomized complete block design so the 290 ANOVA structure includes both a Year and Block within Year random effect. Two treatment 291 structures were considered, i) Genotype / Time looking to assess variety differences and 292 Variety.Time interactions and ii) (Group / Genotype) / Time where Group classifies varieties 293 according to the year of introduction as defined in Figure 1. Variables were transformed, as 294 detailed in Table 2 and Supplementary Table S3, to ensure homogeneity of variance.

295 Multivariate analyses were used to assess variation across all variables. Fibre components 296 were considered as a composition (relative percentage out of 100) and as such were first 297 transformed according to the centred log ratio transformation[30]. To ensure subsequent 298 multivariate analyses focussed on variation between cultivars, variables were adjusted by 299 the Year.Block BLUPs before input into the PCA. PCA was done on the correlation matrix.

300 All statistical analyses were done using Genstat 20th edition.

301 **Table 1**. ANOVA of the treatment effects for Single Kernel Characterisation System (SKCS) measurements; grain kernel weight, diameter and
302 hardness index.

Variable	Cultivar/Time				(Group\Cultivar)/Time							
	Cultivar	Time.Cultivar	Cultivar	Time.Cultivar	Group	Time.Group	Group.Cultivar	Time.Group.Cultivar	Group	Time.Group	Group.Cultivar	Time.Group.Cultivar
		F statistic	F statistic	p-value	p-value	F statistic	F statistic	F statistic	p-value	p-value	p-value	p-value
Kernel weight	52.29	52.07	1.76E-91	0.00E+00	106.41	35.93	49.28	52.96	2.99E-33	3.31E-23	1.92E-87	0.00E+00
Kernel diameter	31.23	3.83	4.67E-70	4.66E-15	115.18	16.32	26.57	3.14	3.59E-35	9.39E-12	2.05E-62	5.31E-11
Hardness index	89.36	21.33	3.16E-115	0.00E+00	504.36	52.26	66.31	19.62	7.22E-84	4.35E-31	2.70E-100	0.00E+00

303

304 **Table 2.** ANOVA of the treatment effects for fibre components and polar metabolite variables.
305

Row number	Variable	Transformation	Figure 3	Cultivar\Time				Group\Cultivar\Time							
				Cultivar	*Time.Cultivar*	*Cultivar*	*Time.Cultivar*	*Group*	*Time.Group*	*Group.Cultivar*	*Time.Group.Cultivar*	*Group*	*Time.Group*	*Group.Cultivar*	*Time.Group.Cultivar*
				F statistic	*F statistic*	*p-value*	*p-value*	*F statistic*	*F statistic*	*F statistic*	*F statistic*	*p-value*	*p-value*	*p-value*	*p-value*
1	Arabinoxylan	Log	A	29.57	2.27	3.63E-67	1.80E-06	212.17	1.06	19.43	2.34	3.75E-52	3.76E-01	1.92E-50	1.15E-06
2	Total beta glucan	Log	B	17.40	1.31	8.84E-48	6.51E-02	36.47	5.71	16.34	1.07	2.05E-14	2.16E-04	1.36E-44	3.49E-01
3	Total amino acid	Log	C	10.82	1.14	7.07E-33	2.25E-01	44.03	1.65	8.98	1.12	8.03E-17	1.63E-01	5.22E-27	2.70E-01
4	Asparagine	Log	D	4.37	1.40	1.67E-12	3.25E-02	27.02	3.34	3.11	1.29	3.18E-11	1.11E-02	1.60E-07	8.48E-02
5	Total carbohydrates	Log	E	31.54	2.80	1.10E-69	2.24E-09	241.95	4.30	19.85	2.72	2.18E-56	2.29E-03	3.40E-51	1.08E-08
6	Total methyl donors	Log	F	30.76	2.25	1.06E-68	2.41E-06	152.41	11.54	24.00	1.73	2.53E-42	1.58E-08	5.33E-58	1.29E-03
7	Total organic acids		G	8.97	1.55	9.27E-28	7.27E-03	16.34	0.97	8.56	1.59	2.40E-07	4.26E-01	8.08E-26	5.94E-03
8	AGP		H	7.18	1.41	2.79E-22	2.88E-02	9.53	5.04	7.05	1.21	1.07E-04	6.59E-04	3.13E-21	1.52E-01

306
307
308

309 **References**

310 1. Mackay, I. et al. Reanalyses of the historical series of UK variety trials to quantify the
311 contributions of genetic and environmental factors to trends and variability in yield
312 over time. *Theor. Appl. Genet.* **122**, 225-238 (2011).
313 2. Pont, C. et al Tracing the ancestry of modern bread wheats. *Nature Genetics* **51**,
314 905-911 (2019).
315 3. Morris, C.E. & Sands, D.C. The breeder's dilemma – yield or nutrition? *Nature*
316 *Biotechnol.* **24**, 1078-1080 (2003).
317 4. Sands, D.C., Morris, C.E., Dratz, E.A. & Pilgeram, A.L. Elevating optimal human
318 nutrition to a central goal of plant breeding and production of plant-based foods. *Plant*
319 *Sci.* **177**, 377-389 (2009).
320 5. Kasarda, D.D. Can an increase in celiac disease be attributed to an increase in the
321 gluten content of wheat as a consequence of wheat breeding? *J. Agric. Food Chem.*
322 **61**, 1155-1159 (2013).
323 6. Shewry, P.R., Pellny, T.K. & Lovegrove, A. Is modern wheat bad for health? *Nature*
324 *Plants* **2**, 1-3(2016).
325 7. Fan, M., Zhao, F.-J., Fairweather-Tait, S.J., Poulton, P.R., Dunham, S.J. & McGrath,
326 S.P. -Set al. Evidence of decreasing mineral density in wheat grain over the last 160
327 years. *Journal of Trace Elements in Medicine and Biology* 22: 315-324 (2008).
328 8. Murphy, K.M., Reeves, P.G. & Jones, S.S. Relationship between yield and mineral
329 nutrient concentrations in historical and modern spring wheat cultivars. *Euphytica*
330 **163,** 381-390 (2008).
331 9. Shewry, P.R., et al. Relationship between the contents of bioactive components in
332 grain and the release dates of wheat lines in the HEALTHGRAIN diversity screen. *J.*
333 *Agric. Food Chem.* **59**, 928-933 (2011)
334 10. Dinelli, G. et al. Determination of phenolic compounds in modern and old varieties of
335 durum wheat using liquid chromatography coupled with time-of-flight mass
336 spectrometry. *Journal Chromatog. A* **1216,** 7229-7240 (2009).
337 11. Heimler, D. et al. Polyphenol content of modern and old varieties of *Triticum*
338 *aestivum* L. and *T. durum* Desf. grains in two years of production. *J. Agric. Food*
339 *Chem.* **58,** 7329-7334 (2010).
340 12. De Santis, MA., Kosik, O., Passmore, D., Flagella, Z., Shewry, P. R., & Lovegrove, A.
341 Comparison of the dietary fibre composition of old and modern durum wheat
342 (*Triticum turgidum* spp. durum) genotypes. *Food Chem.* **244**, 304-310 (2018)

343 13. Johnston, A. E. The Rothamsted classical experiments. In *Long-Term Experiments in*
344 *Agricultural and Ecological Sciences*; Leigh, R. A.; Johnston, A. E., Eds.; CAB
345 International: Wallingford, UK, 1994; pp 9-37.
346 14. Berry, P.M., Kendall, S., Zutterford, Z., Orford, S. & Griffiths, S. Historical analysis of
347 the effects of breeding on the height of winter wheat (*Triticum aestivum*) and
348 consequences for lodging. *Euphytica* **203**, 375-383 (2015).
349 15. Cox, T.S., Shroyer, J.P., Ben-Hui, L.., Sears, R.G., & Martin, T.J. Genetic
350 improvement in agronomic traits of hard red winter wheat cultivars from 1919 to
351 1987. *Crop Sci*. **28**, 756-760 (1988).
352 16. McCaig, T.N. & DePauw, R.M. Breeding hard red spring wheat in western Canada:
353 historical trends in yield and related variables. *Can. J. Pl. Sci.* **75**, 387-393 (1995).
354 17. Steer, T., Thane, C., Stephen, A. & Jebb, S. 'Bread in the diet: consumption and
355 contribution to nutrient intakes of British adults', *Proc. Nutr. Soc.* **67**, E363 (2008).
356 18. Curtis, T.Y. & Halford, N. G. Reducing the acrylamide-forming potential of wheat.
357 *Food and Energy Security* **5**, 153-164 (2016).
358 19. . Corol, D.I., et al. Effects of genotype and environment on the contents of betaine,
359 choline and trigonelline in cereal grains. *J. Agric. Food Chem*. **60**, 5471-81(2012).
360 20. Tryfona, T., et al. Carbohydrate Structural Analysis of Wheat Flour Arabinogalactan
361 Protein. *Carbohydrate Res.* **345**, 2648–2656 (2010).
362 21. Loosveld, A. et al. Structural Variation and Levels of Water-Extractable
363 Arabinogalactan-Peptide in European Wheat Flours. *Cereal Chem.*, **75**, 815–819
364 (1998).
365 22. . Harris, S. et al. Determination of the prebiotic activity of wheat arabinogalactan
366 peptide (AGP) using batch culture fermentation. *Eur. J. Nutrition*. DOI:
367 10.1007/s00394-019-01908-7 (2019).
368 23. Gibson, P. & Shepherd, S. J. Evidence-based dietary management of functional
369 gastrointestinal symptoms: The FODMAP approach. *J. Gastroenterol. Hepatol*. **25**,
370 252-258 (2010).
371 24. Zeisel, S. H. & Blusztajn,J. K. Choline and Human-Nutrition. *Annu. Rev. Nutr*. **14**,
372 269-96 (1994).
373 25. Zheeng X., Levine D., Shen J., Gogarten S.M., Laurie C and Weir B. S. A high
374 performance computing toolset for relatedness and principal component analysis of
375 SNP data. *Bioinfornatics* **28**, 3326-3328, 2012.
376 26. Lovegrove, A. et al. RNA interference suppression of genes in glycosyl transferase
377 families 43 and 47 in wheat starchy endosperm causes large decreases in
378 arabinoxylan content. *Plant Physiol.* **163**: 95–107(2013).

379 27. Ordaz-Ortiz, J. J., Devaux, M. F. & Saulnier, L. Classification of wheat varieties
380 based on structural features of arabinoxylans as revealed by endoxylanase treatment
381 of flour and grain. *J. Agric. Food Chem.* **53,** 8349–8356 (2005).
382 28. Baker, J. M., et al. A metabolomic study of substantial equivalence of field-grown
383 genetically modified wheat. *Plant Biotechnol. J.* **4**, 381-392 (2006).
384 29. Shewry, P.R., Corol, D., Jones, H.D., Beale, M.H. & Ward, J.L. Defining genetic and
385 chemical diversity in wheat grain by 1H-NMR spectroscopy of polar metabolites. *Mol.*
386 *Nutr. Food Res.* **61**, 1600807 (2017).
387 30. Aitchison, J. The Statistical Analysis of Compositional Data. *Monographs on statistics*
388 *and applied probability (Series)* (Chapman and Hall, 1986)
389

390 **Table Legends**

391 **Table 1.** ANOVA of the treatment effects for Single Kernel Characterisation System (SKCS)
392 measurements; grain kernel weight, diameter and hardness index.

393 All treatment effects were tested on 225 residual degrees of freedom. Table shows both the
394 F statistic (representing the size of the effect) and the p-value (the statistical significance of
395 the effect). See also Supplementary Figure S2; box whisker plot of data with group means
396 indicated by black asterisk (✱).

397 **Table 2.** ANOVA of the treatment effects for fibre components and polar metabolite
398 variables.

399

400 Fibre components; arabinoxylan (AX) and total β-glucan; and polar metabolites; total free
401 amino acids, free asparagine, total carbohydrates (mono-, di- and tri-saccharides), total
402 methyl donors (betaine and choline), total organic acids and arabinogalactan peptide (AGP).
403 All treatment effects were tested on 221 residual degrees of freedom. Table shows both the
404 F statistic (representing the size of effect) and the p-value (the statistical significance of the
405 effect). See also Figure 3; box whisker plot of data with group means indicated by black
406 asterisk (✱).

407 **Figure Legends**

408 **Figure 1.** Details of the cultivars used for the study. Three groups are coloured as in the
409 figures.

410 Red, empirical selection (1790-1916); green, early breeding (1935-1972), blue, modern
411 breeding (1980-2012). [1]Selected in Canada from a shipment from Central Europe via
412 Scotland.

413 **Figure 2.** Genomic relationships of the 39 cultivars, illustrated by Principal Component
414 Analysis of markers determined using the Axiom HD Genotyping Array (comprising 819,571
415 SNP markers).

416 Cultivars are coloured to indicate three groups based representing stages in wheat breeding:
417 empirical selection (1790-1916), early breeding (1935-1972) and modern breeding (see
418 Figure 1). Two outliers indicated, labelled (April Bearded, 1884; Apollo, 1986) are discussed
419 in the text.

420 **Figure 3.** Box and whisker plots of the contents of fibre components and polar metabolites
421 in white flour of the 39 cultivars.

422 Cultivars are grouped to represent stages in wheat breeding: red, empirical selection (1790-
423 1916); green, early breeding (1935-1972), blue, modern breeding (1980-2012) (see Figure
424 1)

425 Each point represents the mean of 3 biological replicates. Error bar gives the least
426 significance difference (LSD) between genotype means. A, arabinoxylan (AX); B, β-glucan;
427 C, total free amino acids; D, free asparagine; E, total carbohydrates (mono-, di- and tri-
428 saccharides); F, total methyl donors (betaine and choline); G, total organic acids; H,
429 arabinogalactan peptide (AGP).

430 **Figure 4.** Principle Component Analysis (PCA) of fibre composition (A, B) and polar
431 metabolite composition (C, D) in white flour of the 39 cultivars.

432 Each point represents the mean of 3 biological replicates. Data points in A and C are
433 coloured to indicate three groups of cultivars based representing stages in wheat breeding:
434 red, empirical selection (1790-1916); blue, early breeding (1935-1972); green, modern
435 breeding (see Figure 1). A, scores plot of proportions of oligosaccharides released by
436 digestion of arabinoxylan (AXOS) and β-glucan (G3 and G4 GOS) with endoxylanase and
437 lichenase, respectively; B, loadings plot showing the contributions of AXOS and GOS to the
438 separation in A; C, scores plot of abundances of polar metabolites; D loadings plot showing
439 the contributions of individual metabolites to the separation shown in C.

440 **Acknowledgements**

441 Rothamsted Research receives grant-aided support from the Biotechnology and Biological
442 Sciences Research Council (BBSRC) of the UK and the work forms part of the Designing
443 Future Wheat strategic programme (BB/P016855/1). We thank Dr Ahmed Tawfike for
444 assistance in metabolite quantitation and Mr Gianluca Ruvo for NMR data collection. We are

grateful to Dr Simon Griffiths (John Innes Centre, UK) and Dr Bill Angus (F1 Seed Ltd., Rattlesden, UK) for advice on wheat genotypes.

Author information

Affiliations

Rothamsted Research, Harpenden, Herts, AL52JQ, UK.

Alison Lovegrove, Till Pellny, Kirsty Hassall, Amy Plummer, Abigail Wood, Alice Bellisai, Jane L. Ward, Peter R Shewry

Life Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ

Alexandra Przewieslik-Allen, Amanda J. Burridge

Contributions

AL contributed to study design and writing, supervised fibre analyses. TKP contributed to study design and writing and designed and supervised the field trials. AP and AW carried out fibre analyses. AB carried out NMR analyses. AP-A. and AJB carried out Axiom mapping. JLW carried out metabolite analyses, analysed data, prepared figures and contributed to writing. KLH carried out the statistical analysis. PRS led study design and writing.

Corresponding author

Correspondence to Peter Shewry or Alison Lovegrove

Ethics declaration

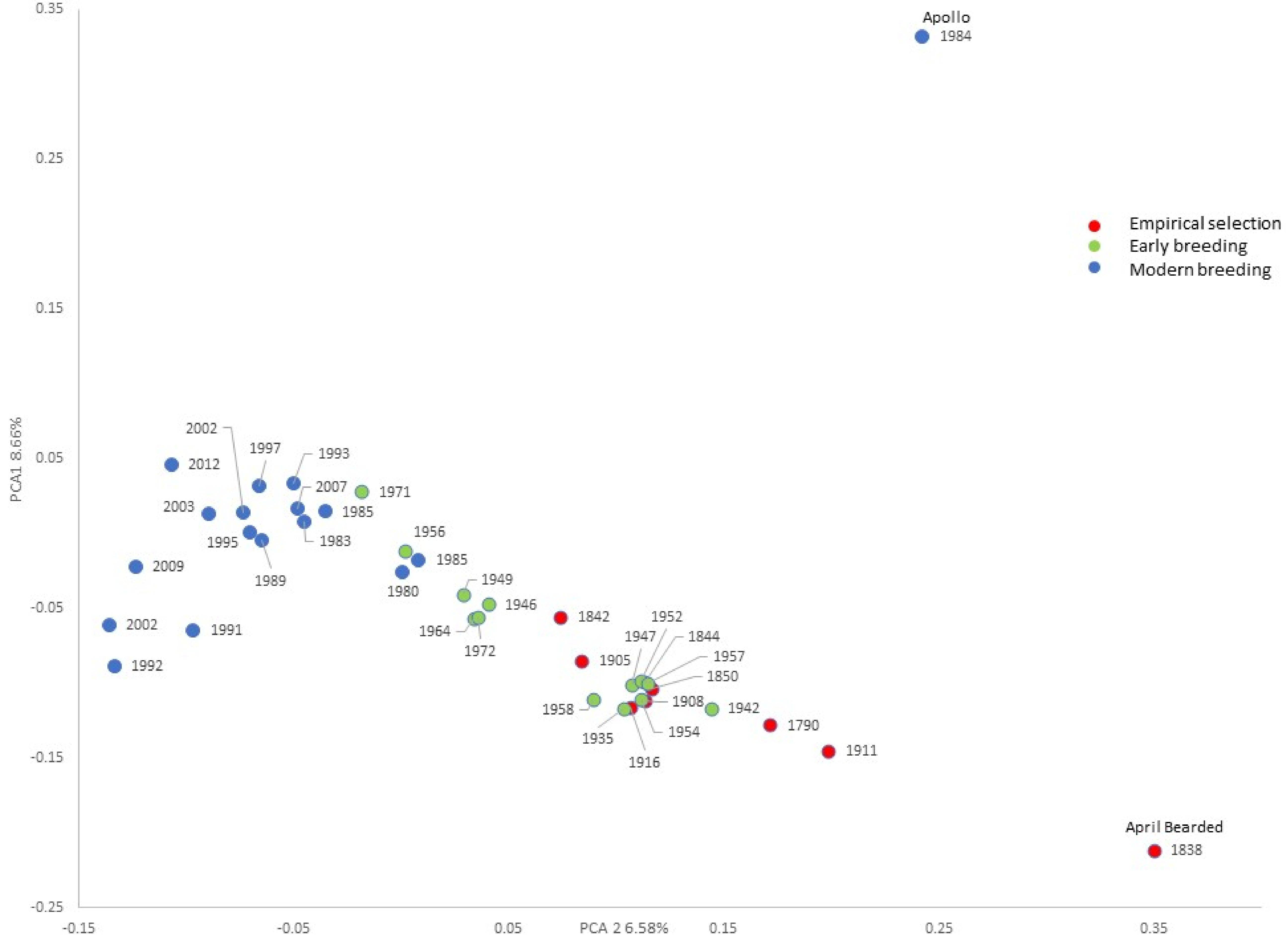
The authors declare no competing interests

Data Availability

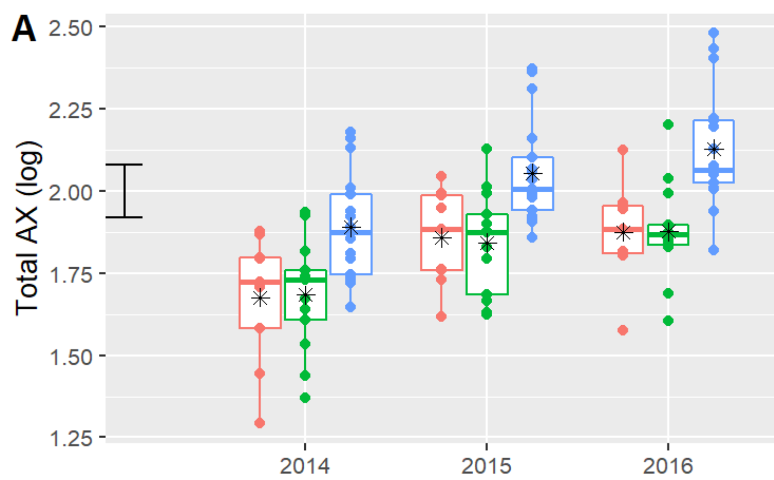
All raw data used in the reported study has been made available in the supplementary files submitted.

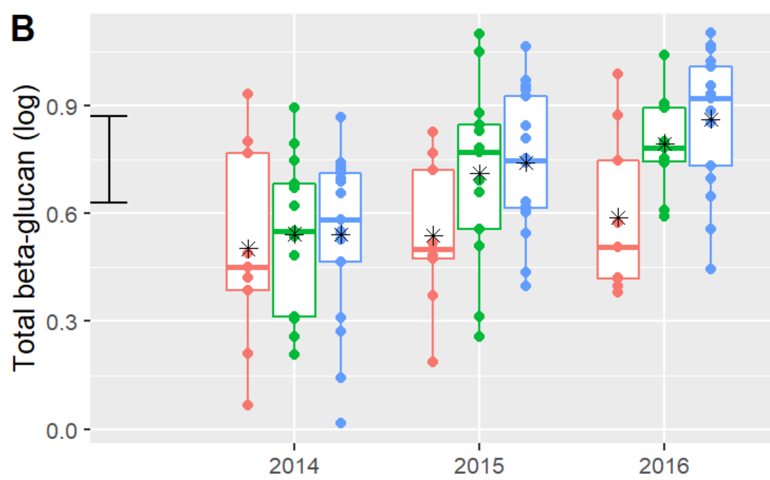
Details of cultivars used in the study

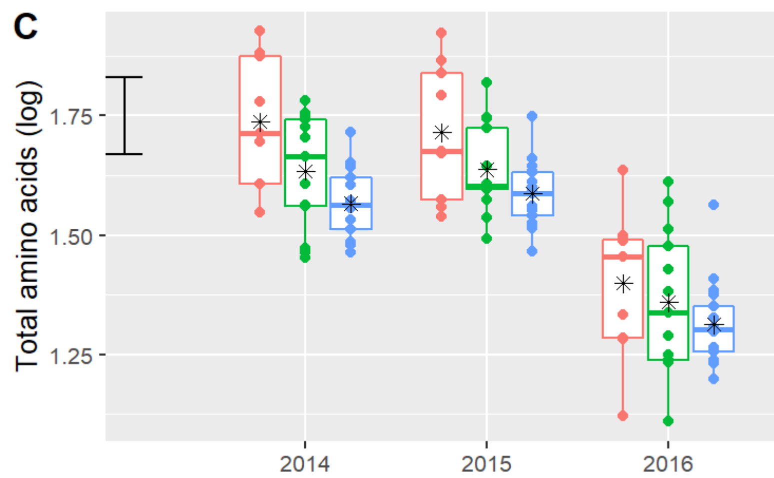
Date	Cultivar	Allele RhtB1	Allele RhtD1	Rht phenotype	Country of origin (if not UK)
1790	Chidham White Chaff				
1838	April Bearded				
1842	Red Fife				Canada[1]
1844	Browick				
1850	Red Lammas				
1905	Red Standard				
1908	Little Joss				
1911	Squareheads Master				
1916	Yeoman				
1935	Holdfast				
1942	Garton Sixty				
1946	Cappelle Desprez				France
1947	Victor				
1949	Flanders				Belgium
1952	Steadfast				
1954	Master Piece				
1956	Viking				
1957	Dominator				
1958	Milfast				
1964	Maris Widgeon				
1971	Maris Huntsman				
1972	Maris Ploughman				
1980	Avalon	Rht-B1a	Rht-D1b	Rht2	
1983	Galahad	Rht-B1a	Rht-D1b	Rht2	
1984	Apollo	Rht-B1a	Rht-D1a	WT	Germany
1985	Mercia	Rht-B1a	Rht-D1b	Rht2	
1985	Brimstone	Rht-B1a	Rht-D1b	Rht2	
1989	Hereward	Rht-B1a	Rht-D1b	Rht2	
1991	Spark	Rht-B1a	Rht-D1a	WT	
1992	Cadenza	Rht-B1a	Rht-D1a	WT	
1993	Consort	Rht-B1a	Rht-D1b	Rht2	
1995	Flame	Rht-B1a	Rht-D1b	Rht2	
1997	Malacca	Rht-B1a	Rht-D1b	Rht2	
2002	Solstice	Rht-B1a	Rht-D1b	Rht2	
2002	Xi 19	Rht-B1a	Rht-D1b	Rht2	
2003	Robigus	Rht-B1b	Rht-D1a	Rht1	
2007	Einstein	Rht-B1a	Rht-D1b	Rht2	
2009	Gallant	Rht-B1a	Rht-D1b	Rht2	
2012	Crusoe	Rht-B1a	Rht-D1b	Rht2	

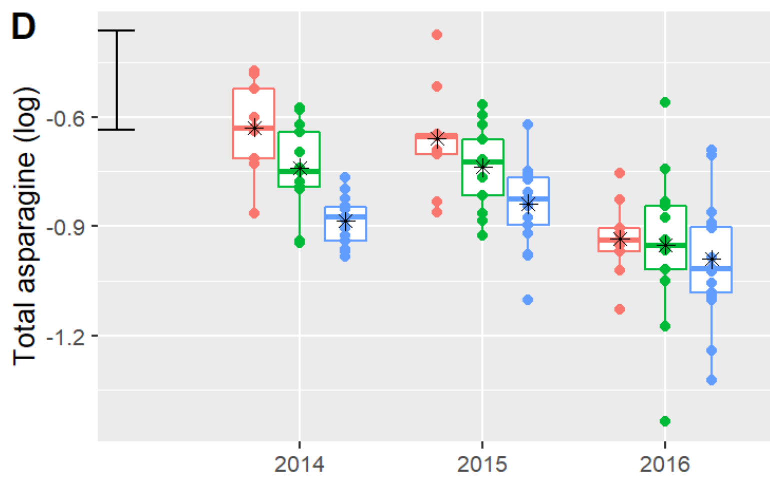


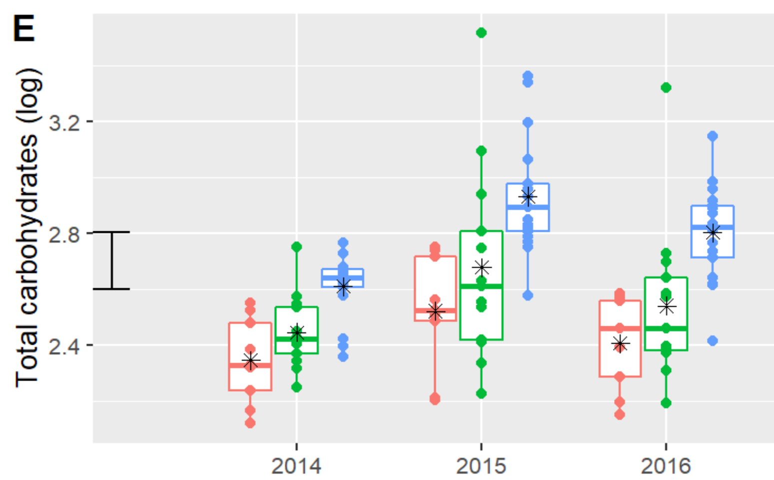
Apollo
1984
Empirical selection
Early breeding
Modern breeding
PCA1 8.66%
PCA 2 6.58%
April Bearded
1838
2012
2002
1997
1993
2007
1971
2003
1985
1995
1983
1989
2009
1956
1985
1980
1949
1946
2002
1991
1992
1842
1964
1972
1952
1947
1844
1905
1957
1850
1958
1908
1942
1790
1935
1954
1916
1911
0.35
0.25
0.15
0.05
-0.05
-0.15
-0.25
-0.15
-0.05
0.05
0.15
0.25
0.35

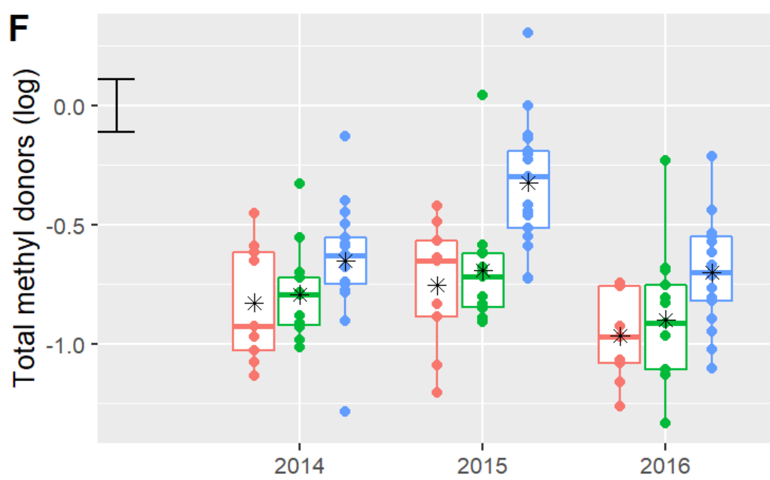

A
2.50
2.25
2.00
1.75
1.50
1.25
Total AX (log)
2014
2015
2016

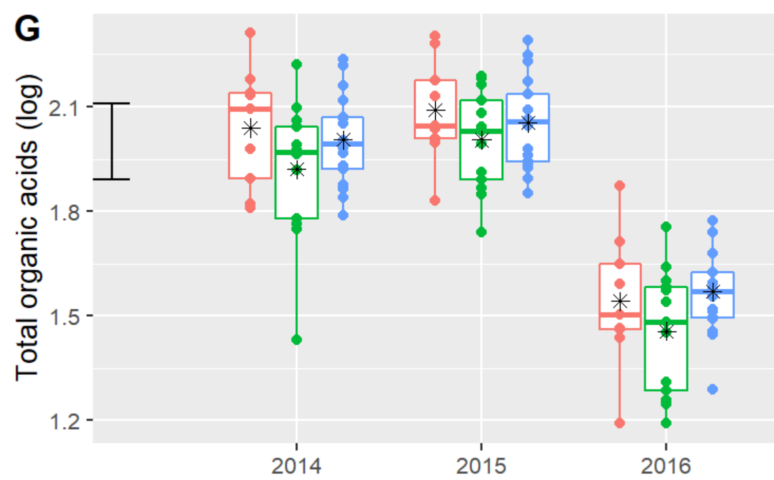

B
0.9
0.6
0.3
0.0
Total beta-glucan (log)
2014
2015
2016

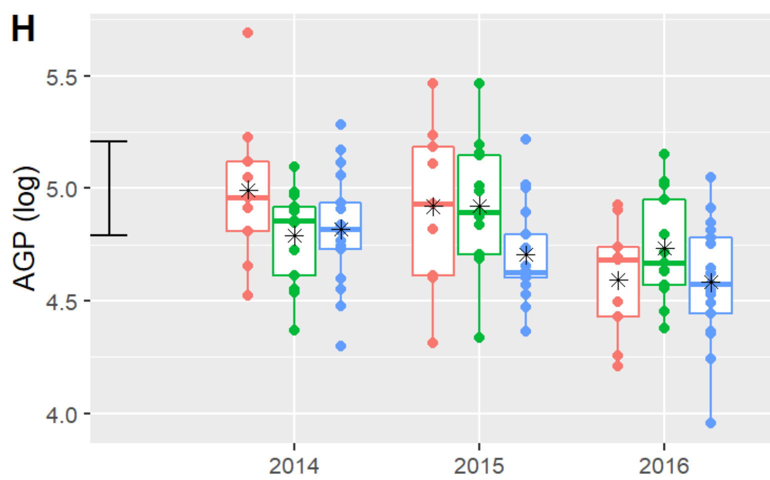

C
1.75
1.50
1.25
Total amino acids (log)
2014
2015
2016

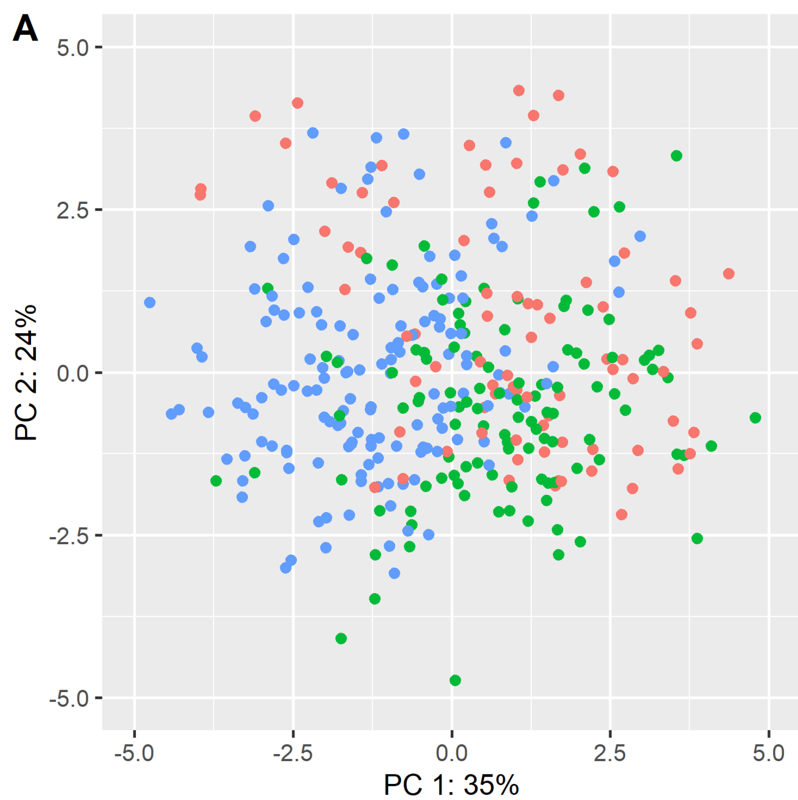

D
-0.6
-0.9
-1.2
Total asparagine (log)
2014
2015
2016

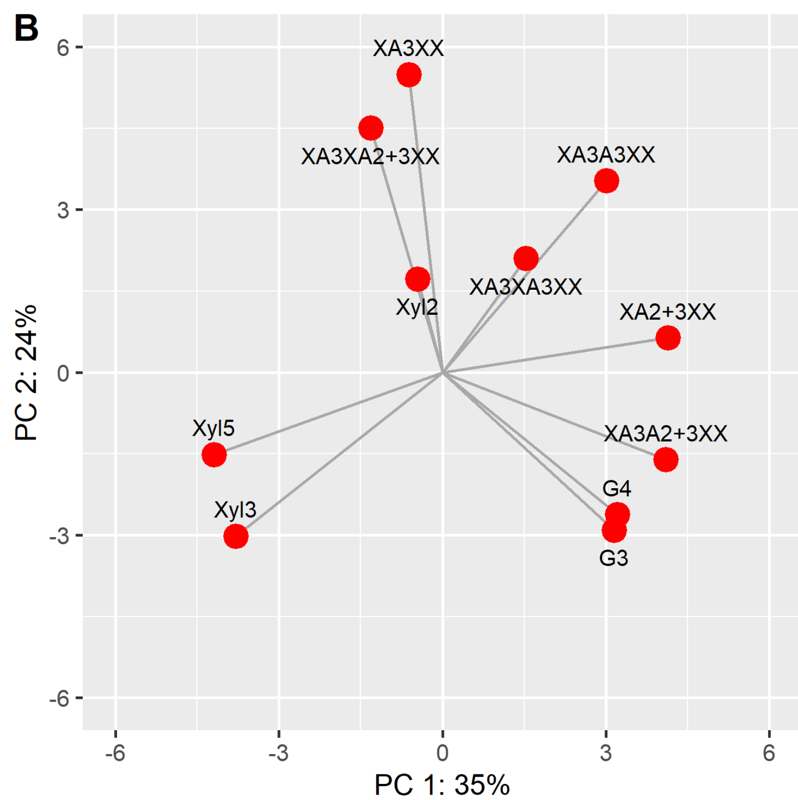

E
3.2
2.8
2.4
Total carbohydrates (log)
2014
2015
2016

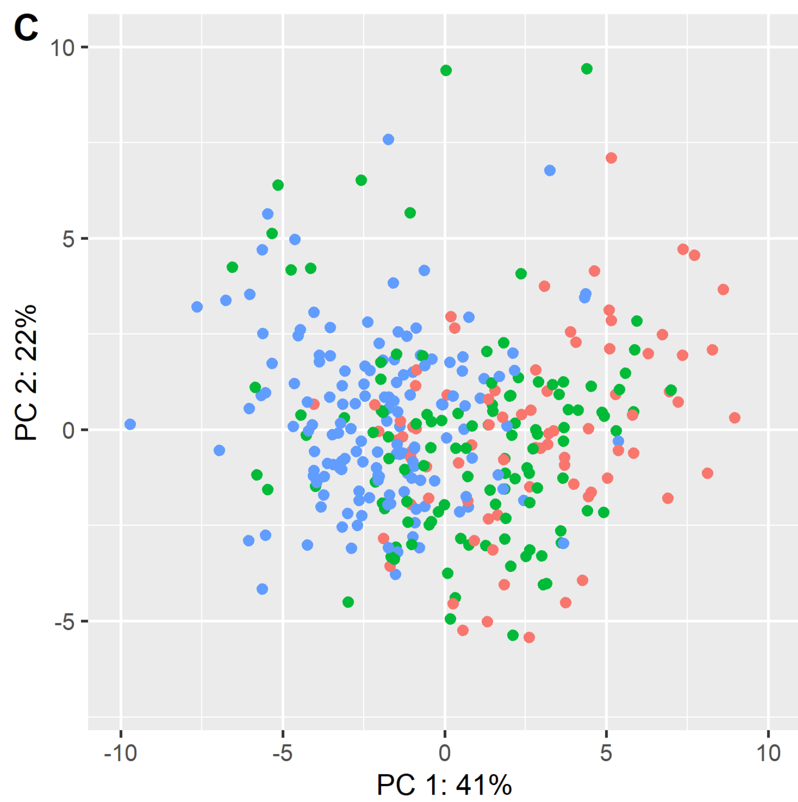

F
0.0
-0.5
-1.0
Total methyl donors (log)
2014
2015
2016

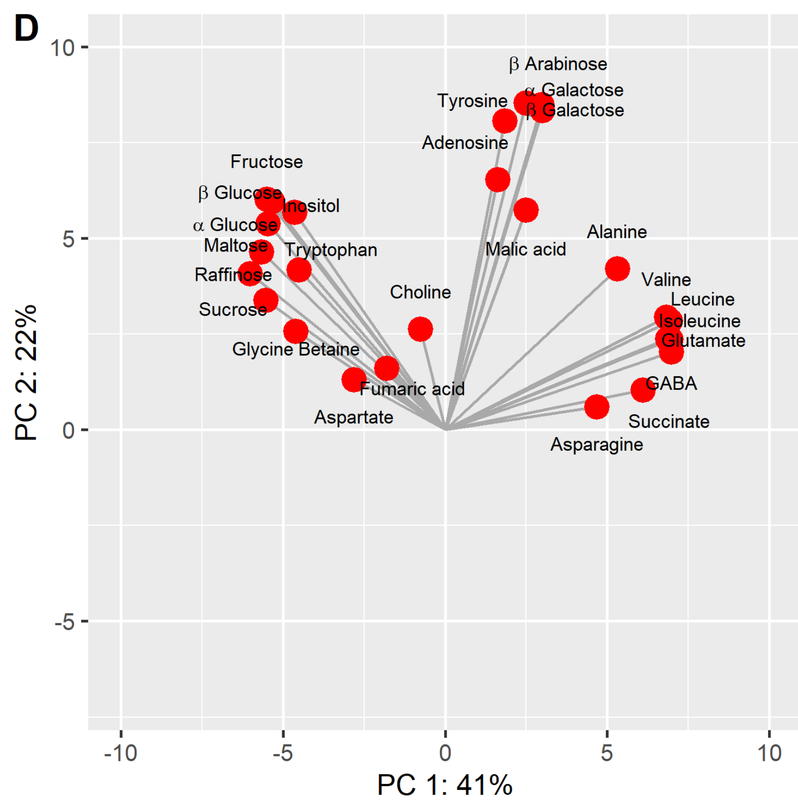

G
2.1
1.8
1.5
1.2
Total organic acids (log)
2014
2015
2016


H
5.5
5.0
4.5
4.0
AGP (log)
2014
2015
2016


A
PC 2: 24%
PC 1: 35%
5.0
2.5
0.0
-2.5
-5.0
-5.0
-2.5
0.0
2.5
5.0


B
PC 2: 24%
PC 1: 35%
XA3XX
XA3XA2+3XX
XA3A3XX
XA3XA3XX
Xyl2
XA2+3XX
Xyl5
XA3A2+3XX
G4
Xyl3
G3
6
3
0
-3
-6
-6
-3
0
3
6


C
PC 2: 22%
PC 1: 41%
10
5
0
-5
-10
-5
0
5
10


D
PC 2: 22%
PC 1: 41%
β Arabinose
α Galactose
Tyrosine
β Galactose
Adenosine
Fructose
β Glucose
Inositol
α Glucose
Maltose
Tryptophan
Malic acid
Alanine
Raffinose
Valine
Choline
Leucine
Sucrose
Isoleucine
Glutamate
Glycine Betaine
GABA
Fumaric acid
Aspartate
Succinate
Asparagine
10
5
0
-5
-10
-5
0
5
10