ORIGINAL ARTICLE

Identification of eight QTL controlling multiple yield components in a German multi-parental wheat population, including *Rht24*, *WAPO-A1*, *WAPO-B1* and genetic loci on chromosomes 5A and 6A

Beatrice Corsi[1] · Lia Obinu[2] · Camila M. Zanella[1] · Saverio Cutrupi[3] · Rob Day[1] · Manuel Geyer[4] · Morten Lillemo[5] · Min Lin[5] · Lorenzo Mazza[3] · Lawrence Percival-Alwyn[1] · Melanie Stadlmeier[4,6] · Volker Mohler[4] · Lorenz Hartl[4] · James Cockram[1]

Abstract

Key message **Quantitative trait locus (QTL) mapping of 15 yield component traits in a German multi-founder population identified eight QTL each controlling ≥2 phenotypes, including the genetic loci *Rht24*, *WAPO-A1* and *WAPO-B1*.**

Abstract Grain yield in wheat (*Triticum aestivum* L.) is a polygenic trait representing the culmination of many developmental processes and their interactions with the environment. Toward maintaining genetic gains in yield potential, 'reductionist approaches' are commonly undertaken by which the genetic control of yield components, that collectively determine yield, are established. Here we use an eight-founder German multi-parental wheat population to investigate the genetic control and phenotypic trade-offs between 15 yield components. Increased grains per ear was significantly positively correlated with the number of fertile spikelets per ear and negatively correlated with the number of infertile spikelets. However, as increased grain number and fertile spikelet number per ear were significantly negatively correlated with thousand grain weight, sink strength limitations were evident. Genetic mapping identified 34 replicated quantitative trait loci (QTL) at two or more test environments, of which 24 resolved into eight loci each controlling two or more traits—termed here 'multi-trait QTL' (MT-QTL). These included MT-QTL associated with previously cloned genes controlling semi-dwarf plant stature, and with the genetic locus *Reduced height 24* (*Rht24*) that further modulates plant height. Additionally, MT-QTL controlling spikelet number traits were located to chromosome 7A encompassing the gene *WHEAT ORTHOLOG OF APO1* (*WAPO-A1*), and to its homoeologous location on chromosome 7B containing *WAPO-B1*. The genetic loci identified in this study, particularly those that potentially control multiple yield components, provide future opportunities for the targeted investigation of their underlying genes, gene networks and phenotypic trade-offs, in order to underpin further genetic gains in yield.

* James Cockram
  james.cockram@niab.com

[1] NIAB, 93 Lawrence Weaver Road, Cambridge CB3 0LE, UK

[2] Department of Agriculture, University of Sassari, Viale Italia, 07100 Sassari, Italy

[3] University of Messina, Messina, Italy

[4] Bavarian State Research Center for Agriculture, Institute for Crop Science and Plant Breeding, 85354 Freising, Germany

[5] Norwegian University of Life Sciences (NMBU), P.O. Box 5003, NO-1432 Ås, Norway

[6] Present Address: Saatzucht Donau GesmbH and Co KG, Mendelweg 1, 4981 Reichersberg, Austria

Introduction

Bread wheat (*Triticum aestivum* L.) is one of the world's most widely cultivated crops, accounting for more than 20% of human calorific intake (FAO 2017). Accordingly, wheat production is a key component underpinning global food security. Wheat grains develop within the inflorescence, more

commonly termed the 'spike' or 'ear', located at the apex of a long stem called a tiller. The wheat ear is composed of a central rachis along which branching structures called the rachilla each bear multiple florets that form the spikelet. Each ear typically contains ~20 spikelets, which are arranged sequentially on alternate sides of the rachis axis with a terminal spikelet at the apex (Murai et al. 2002). The rachilla meristem is indeterminant, and around 4–5 florets are arranged sequentially along alternating sides of the rachilla axis, with the florets referred to from the spikelet base to tip as position G1, G2, G3, etc. Ultimately, grain yield represents the culmination of many genetic, developmental, and environmental interactions with the coordination of plant growth and development across the growing season, and is under complex genetic control—although it can often show high heritability (e.g. Muqaddasi et al. 2019). Yield can be largely partitioned into three major components: grain weight/size and shape, grain number per spike and spike number per unit area (Gegas et al. 2010). Evidence suggests that improvements in grain number per ear (itself a combination of the number of spikelets, and floret fertility) have played an important role in the increased yields attained via modern wheat breeding (e.g. Calderini et al. 1995; Würschum et al. 2018). Similarly, grain size has also increased, with for example a recent study of Chinese wheat varieties indicating that breeding has resulted in increased grain size predominantly at positions G1, G2 and G3 (Feng et al. 2018). These major yield components can be further subdivided into constituent traits, e.g. grain size = grain morphometric traits, carpel (unfertilized female organ) size, rate of grain filling, cell expansion rate, and cell number (Brinton and Uauy 2019). They are also influenced by other factors such as accumulation and transport of photosynthetic products, flag leaf size and angle, plant height, overall biomass and flowering time, and the interaction of all these factors with the environment. Furthermore, wheat processing traits that affect milling performance (e.g. grain shape, size, density and uniformity) are also critical for flour yield. The inter-dependent roles of the 'source' (e.g. leaves) and 'sink' (e.g. grain) tissues in building yield across the lifecycle of a wheat plant are complex, and trade-offs between yield components are the norm. For example, two varieties with the same overall grain yield could be a result of more grains per ear in one variety, or more tillers and less grain per ear in another. Similarly, the strong negative correlation between grain yield and protein content is a well-recognized trade-off in wheat, and other crops (Simmonds 1995; Mosleth et al. 2020). Such compensatory phenotypic effects, along with the large, complex polyploid nature of the 17 Gb wheat genome (IWGSC 2018)—within which homoeologous gene copies buffer the effects of mutation at any single gene—make the physiological and genetic study of yield and yield components potentially more challenging than in diploids (Brinton and Uauy 2019). For example, numerous natural genetic variants controlling yield components

have been map-based cloned in the diploid cereal crop rice (*Oryza sativa* L.) (for a review of rice grain size genes, see Li et al. 2018a), and these tend to account for ~10–40% of the phenotypic variation. In contrast, while exceptions do exist, most notably a locus on chromosome 7A controlling spikelet number per ear (Muqaddasi et al. 2019; Kuzay et al. 2019; Voss-Fels et al. 2019), natural variation at wheat genetic loci controlling yield and yield component traits in bread wheat typically control just ~2–10% of the variation (reviewed by Brinton and Uauy, 2019). This observation is likely due to the compensatory effects of homoeologous genes in polyploid wheat. However, this also provides opportunities for molecular-assisted improvement in hexaploid wheat if characterized alleles at homoeologous genetic loci or genes can be combined in a single genetic background to overcome functional redundancy. One route to achieve this aim is to identify quantitative trait loci (QTL) controlling a trait of interest, and ultimately to characterize the genes and gene variants that underly them. Toward this aim, numerous studies investigating the genetics and/or physiology of grain yield (e.g. Rustgi et al. 2013; Garcia et al. 2019), grain size (e.g. Gegas et al. 2010; Xie et al. 2015; Brinton et al. 2017), carpel size (Xi et al. 2015) and grain three-dimensional shape (Hughes et al. 2019) have been undertaken, as well as analyses of key traits which impact yield such as flowering time (e.g. Griffiths et al. 2009a), height (e.g. Griffiths et al. 2009b), flag leaf size (e.g. Xue et al. 2013; Ma et al. 2020), stomatal conductance and photosynthetic performance (Faralli et al. 2019). However, just two wheat yield component QTL have been cloned or fine-mapped to the genic level to date. The first, *Grain Number Increase 1* (*GNI1*), was identified as a QTL in a cross between a tetraploid durum wheat (*T. durum*) cultivar and a *T. durum* line which contained an introgression from wild tetraploid emmer wheat (*T. turgidum* ssp. *dicoccoides*) (Sakuma et al. 2019). The underlying gene, *GNI-A1*, encodes a homeodomain leucine zipper class I (HD-Zip 1) transcription factor orthologous to the barley inflorescence architecture gene *Six-rowed spike 1* (*VRS1*) that controls the fertility of lateral florets on the barley spikelet (Komatsuda et al. 2007). *GNI-A1* acts to repress floral fertility in durum wheat, with the reduced function allele resulting in increased numbers of fertile florets per spikelet. The *GNI-A1* reduced function allele increases grain weight at spikelets located predominantly at the basal and central regions of the ear (Sakuma et al. 2019; Golan et al. 2019), as well as at specific grain position within the spikelet (G1, G2, G3). Subsequent resequencing of *GNI-A1* in a panel of European bread wheat found a similar reduced function allele to be present in around a third of the varieties surveyed and was associated with higher grain number per ear (Sakuma et al. 2019). More recently, fine-mapping of a QTL on chromosome 7A controlling spikelet number per ear in bread wheat has resulted in the identification of *WHEAT ORTHOLG OF APO1* (*WAPO1*) as the likely underlying gene (Kuzay et al. 2019; Muqaddasi et al.

2019). *WAPO1* underlies a wheat yield component QTL of unusually large effect, explaining up to 23% of the phenotypic variation (Kuzay et al. 2019; Muqaddasi et al. 2019; Voss-Fels et al. 2019). It encodes an F-box protein orthologous to the rice gene *ABBERANT PANICLE ORGANISATION 1* (*APO1*) which controls rice spikelet number (Ikeda et al. 2007), and is homologous to *UNUSUAL FLORAL ORGANS* (*UFO*) that regulates arabidopsis floral organ identity (Wilkinson et al. 1995; Samach et al. 1999).

To date, most studies of yield and yield components have been undertaken in either bi-parental or association mapping populations, each of which has their own inherent pros and cons (reviewed by Cockram and Mackay 2018). For example, bi-parental populations are limited in the genetic diversity that the two founders can capture, while the need to account for genetic population substructure can reduce the power to detect QTL in association mapping panels. The relatively recent development of multi-founder populations in many plant species (reviewed by Scott et al. 2020a) provides alternative population types with which to investigate the genetics of target traits. Here we use an eight founder multiparent advanced generation inter-cross (MAGIC) population: the Bavarian MAGIC winter wheat population (BMW-pop) (Stadlmeier et al. 2018). The population consists of 394 $F_{6:8}$ recombinant inbred lines derived from eight founders: Event, BAYP4535, Ambition, Firl3565, Format, Potenzial, Bussard and Julius. With the exception of the Danish variety 'Ambition', the founders represent breeding lines and released varieties developed by German wheat breeding companies, and the population has been estimated to capture 71.7% of the allelic diversity available in the German wheat breeding gene pool (Stadlmeier et al. 2018). Although the BMWpop has previously been used to investigate the genetics of wheat disease resistance (Stadlmeier et al. 2019; Lin et al. 2020), it has yet been used to study yield components. With the aim of identifying yield component QTL relevant to north-western European wheat cultivation, we grew the BMWpop in three trials across two countries and phenotyped 15 traits. Subsequent genetic analysis resulted in the identification of 34 replicated QTL, of which 24 were found to resolve into eight genetic loci each controlling two or more traits.

Methods

Wheat germplasm and field trials

Details of the characteristics of the BMWpop founders and population are described by Stadlmeier et al. (2018). Briefly, it consists of 394 recombinant inbred lines created from eight winter wheat varieties using a simplified mating design, which included an additional eight-way inter-cross step to give the same results in terms of the levels of missing founder assignments and number of recombination events as a MAGIC design with all 210 possible four-way crosses. The BMWpop genetic linkage map consists of 5,435 single nucleotide polymorphism (SNP) markers distributed over 2,804 loci and spanning 5,230 cM (Stadlmeier et al. 2018). Here, all 394 lines of the BMWpop were grown with two replicate plots per line in three trials: two trials in the United Kingdom (UK, in years 2017 and 2018), and one in Germany in 2018. Trials were sown in the autumn and grown to full maturity the following summer and were cultivated under standard local agronomic practices, but without the application of fungicides and growth regulators in the UK. The two UK field trials were conducted at NIAB, Cambridge, United Kingdom (UK) (Lat. 52.242931, Long. 0.097802). The trials consisted of 394 recombinant inbred lines in two replications each, and the eight founders in three replications each, plus three additional controls (the cultivars Arina, Jenga and Tarso) in four replications each. Trial design was undertaken in R (R Core Team, 2015) using the package Blocks Design v2.8. Each trial was arranged in two randomized, complete replicates, each of 13 blocks for a total of 836 plots, with each plot consisting of two 1-m rows. The German trial was conducted at Frankendorf (Lat. 48.348091, Long. 11.977043) and consisted of the 394 recombinant inbred lines and founders in two replications each, with the exception of the founder Julius which was replicated nine times, as well as the check variety RGT Reform in seven replications. The trial was composed of 860 plots in an alpha lattice design with two replications and 43 incomplete blocks per replication. Three sub blocks were built based on plant height. Randomization of the trial was performed using the R package Agricolae version 1.3–2 (de Mendiburu 2009). Plot sizes were 1.5 m x 3 m. Subsequently in this manuscript, where numbers are used to indicate BMWpop founders, these are listed as: (1) Event, (2) BAYP4535, (3) Ambition, (4) Firl3565, (5) Format, (6) Potenzial, (7) Bussard, (8) Julius.

Phenotyping and trial analyses

The trait 'plant height' was measured in cm as the height from ground to the top of the spike, excluding awns or scurs. 'Flowering time' was scored the date the majority of plants within a plot had fully emerged ears (Zadoks growth stage GS59), and recorded as days from the 1st May. From each plot, ten representative ears were collected at plant maturity for subsequent phenotypic analysis: 'ear length' (in mm, measured from the base of the first spikelet to the apex of the last spikelet at the top of the ear, excluding awns or scurs), 'ear weight' (g), 'total number of spikelets', 'number of fertile spikelets,' and 'number of infertile spikelets'. Subsequently, ear samples from each plot were threshed, cleaned

of chaff, and used for grain phenotyping using digital images taken with a MARVIN Grain Analyser (www.gta-sensorik. com). Analysis of the resulting images using MARVIN 5.0 software (GTA Sensorik GmbH), combined with the knowledge of the number of ears from which the grain originated, resulted in the calculation of the following phenotypes: 'number of seeds per ear', 'seed length', 'seed width,' and 'seed area'. After imaging, grain was weighed (g), allowing 'thousand grain weight' (g) to be determined using the MARVIN software. The following additional derived phenotypes were also calculated: 'function form density' (thousand grain weight/seed area), 'seed length/seed width ratio', and 'seed weight/ear weight ratio'. Finally, the most prominent fungal diseases present in the UK17 (septoria tritici blotch) and UK18 (yellow rust) trials were scored on the first four leaves at the plot level using a percentage infection scale, following the guide in Supplementary Table 1.

Statistical analysis and genetic mapping

Summary statistics (mean, median, standard deviation and variance) were calculated using GenStat 19th edition (VSN International). Best linear unbiased estimates (BLUEs) were calculated using a linear mixed approach in REML using GenStat. For each trait in each test environment, histograms of the BLUEs were created using the package ggplot2 (Wickham 2016) in R Studio (RStudio Team 2020). Pearson's correlation coefficients and paired Wilcoxon signed-rank test were carried out in R Studio using the package Hmisc (Harrell 2019) and plotted using the R package corrplot (Wei and Simko 2017). Heritability was estimated as $1 - Att / (2 \gamma g^2)$, where γg^2 is the variance component of the term divided by the residual variance, and Att is the average variance for differences between the effects of the term divided by the residual variance (Cullis et al. 2006; Piepho and Möhring 2007).

MAGIC QTL mapping was carried out using the subset of 2,804 SNP markers representing unique positions on the BMWpop genetic map, as described by Stadlmeier et al. (2018). Four QTL analysis approaches were used: (1) SMA (identity-by-state single marker analysis): single marker analysis using the SNP bi-allelic classes using R/lme4 (Bates et al. 2015). (2) IBD (identity by descent): regression against haplotype probability estimates calculated using the 'mpprob' function in R/mpMap (Huang and George 2011) implemented in R/qtl (Broman et al. 2003) with a threshold of 0.5. (3) IM (interval mapping): conducted in R/mpMap using R/mpMap haplotype probability estimates. (4) CIM (composite interval mapping): conducted in R/mpMap with 5 (CIM-cov5) or 10 (CIM-cov10) covariates using R/mpMap haplotype probability estimates. For IBS and IBD analyses, multiple-test correction was carried out using the p.adjust function in R, with a threshold of $p < 0.05$. For IM/CIM

an empirical p threshold of 0.05 was determined for QTL analyses by conducting 100 simulations, using the sim.sigthr function in R. This value, together with a window size of 100 markers, was used to determine QTL peaks using "find. qtl". All detected QTL were simultaneously fitted in a full model using the function 'fit.qtl, and QTL retained where $p < 0.05$ and percentage variation explained $> 1\%$ in the fitted model. The full model including all detected QTL was fitted to estimate the overall phenotypic variance explained (R^2). IM was used to call QTL, with additional detection using CIM-cov5, CIM-cov10. IBD and IBS were undertaken to further confirm IM/CIM QTL. Significance values and percentage variation explained for all QTL reported in the manuscript are derived from IM. Flanking markers were defined by CIM-cov10 or CIM-cov5 when QTL were detected by both IM and CIM; otherwise, intervals were defined by IM. QTL detected in IM analysis were named following the recommendations for gene symbolization in wheat (McIntosh et al. 2013).

Bioinformatics, DNA sequencing, and haplotype analysis

The position of SNP markers on the wheat reference genome RefSeq v1.0 (IWGSC 2018) was obtained from the T3 wheat website (https://triticeaetoolbox.org/wheat/). Where not available, or where the retrieved chromosome assignation differed to that of the BMWpop genetic map (Stadlmeier et al. 2018), the SNP DNA flanking sequenced was used as a query for BLASTn (Altschul et al. 1990) searches of the RefSeq v1.0 wheat genome assembly using the BLAST function at EnsemblPlants release 49 (https://plants.ensembl.org/; Howe et al. 2019) using the default search parameters, and the position of the SNP on the correct chromosome determined manually. To determine whether QTL mapped to regions of low genetic recombination, for each chromosome, marker genetic map locations (Stadlmeier et al. 2018) were plotted against their RefSeq v1.0 physical map positions, where available, using Microsoft Excel. *WAPO-A1* and *WAPO-B1* genomic sequence for the BMWpop founders was generated by direct Sanger sequencing of PCR products using the homoeologue-specific primers listed in Supplementary Table 2, following the protocols described by Cockram et al. (2007). The sequences are available in GenBank under accession numbers MW366865 to MW366879. *WAPO-A1* and *WAPO-B1* sequences for an additional 16 hexaploid wheat lines were extracted from published genome assemblies (IWGSC 2018; Walkowiak et al. 2020). DNA and protein alignments were undertaken using Clustal Omega (Madeira et al. 2019) and visualized using GeneDoc v2.7 (Nicholas et al. 1997). *WAPO* haplotypes were determined manually using sequence alignments, with the positions of DNA variants listed relative to start

codon in the relevant cv. Chinese Spring gene model. Protein domains were identified using Pfam 33.1 (El-Gebali et al. 2019). Visualization of amino acid sequence conservation at specific regions was undertaken in WebLogo (Crooks et al. 2004) using the predicted proteins of plant species with sequenced genomes identified as 'orthologous' in Ensembl Plants release 49, and for which amino acid sequence similarity with WAPO-A1 or WAPO-B1 was > 20% (listed in Supplementary Table 3). For the traits that co-located at the chromosome 7A locus controlling spikelet number phenotypes, haplotype analysis was conducted using the meta-analysis BLUEs, from which means for each trait were calculated and the RILs grouped according to marker *wsnp_JD_c20555_18262317* as a surrogate for *WAPO-A1* haplotype. Haplotype analysis for the 7B loci controlling spikelet number traits spanning *WAPO-B1* were conducted using the peak genetic marker identified in the meta-analysis for both 'total number of spikelets' and 'number of fertile spikelets' (*AX.94949800*) and the two flanking markers immediately flanking this SNP in the BMWpop genetic map (*Tdurum_contig81911_179* and *Kukri_c12901_706*). To account for the strong effect of the 7A locus on spikelet number traits, we used SNP *wsnp_JD_c20555_18262317* as a co-factor to generate the BLUEs used for the haplotype analysis of the 7B QTL. The significance of the difference between the means for each haplotype at the 7A or 7B locus was determined by Wilcoxon test at $p < 0.05$ using the R/ggpubr package (Kassambara 2020). Gene expression data was sourced from Li et al. (2018b) and Ramírez-González

et al. (2018) following the methods outlined in Supplementary Text 1, and heatmaps generated using the heatmap.2 function in the gplots package (Warnes et al. 2012) in R.

Results

Phenotyping, transgressive segregation, and heritabilities

Three replicated field trials were undertaken using the BMWpop: two in the UK in 2017 and 2018 (UK17, UK18) and one in Germany in 2018 (DE18). Across all 394 RILs and eight founders, 15 phenotypes were collected (listed in Table 1): 13 yield components, as well as 'plant height' and 'flowering time'. BLUEs were calculated for all traits and are listed in Supplementary Table 4. Phenotypic data for all traits were found to be normally distributed, and so did not require further transformation. Transgressive segregation was observed in both the positive and negative direction for all traits in all 45 year/location combinations analyzed (Supplementary Fig. 1). This was especially notable for 'number of infertile spikelets per ear', 'total number of spikelets per ear', 'seed length' and 'flowering time'. Heritabilities varied depending on phenotype and year (Table 1), ranging from 0.42 ('seed weight/ear weight ratio' in trial UK17) to 0.95 ('seed length–width ratio' in DE18). Mean heritability for seed traits (SA, SWI, SL, FFD, SL.SWI, TGW; mean $h^2 = 0.80$) were higher than those for ear architecture (EL,

Table 1 Trait descriptions, trait heritabilities (h^2), and trait correlations between trials. Listed are the 15 traits measured in the BMWpop over three trials, conducted in the United Kingdom in 2017 (UK17) and 2018 (UK18), and in Germany in 2018 (DE18)

Trait code	Trait name	UK17 h^2	UK18 h^2	DE18 h^2	Max. h^2 difference	Corr. UK17 vs. UK18	Corr. UK17 vs. DE18	Corr. UK18 vs. DE18
EL	Ear length (cm)	0.77	0.75	0.86	0.11	0.71	0.74	0.80
EW	Ear weight (g)	0.60	0.69	0.80	0.20	0.65	0.29	0.22
NFSP	No. fertile spikelets per ear	0.77	0.66	0.85	0.19	0.71	0.74	0.70
NISP	No. infertile spikelets per ear	0.57	0.57	0.78	0.21	0.35	0.43	0.14
totNSP	Total no. spikelets per ear	0.85	0.70	0.81	0.15	0.76	0.83	0.76
NS.NE	Number of seeds per ear	0.65	0.46	0.85	0.39	0.42	0.45	0.26
WS.EW	Seed weight/ear weight ratio	0.42	0.84	0.62	0.42	0.11	0.16	0.04
SA	Seed area (mm^2)	0.77	0.85	0.91	0.14	0.67	0.53	0.47
SWI	Seed width (mm)	0.71	0.83	0.84	0.13	0.6	0.45	0.36
SL	Seed length (mm)	0.84	0.89	0.69	0.20	0.73	0.60	0.49
FFD	Factor form density	0.72	0.77	0.75	0.05	0.61	0.40	0.33
SL.SWI	Seed length/seed width ratio	0.78	0.89	0.95	0.17	0.66	0.66	0.53
FT	Flowering time (days)	0.82	0.63	0.88	0.25	0.73	0.76	0.74
HT	Plant height (cm)	0.83	0.77	0.93	0.16	0.68	0.60	0.74
TGW	Thousand grain weight (g)	0.74	0.61	0.93	0.32	0.66	0.49	0.46
Mean		0.72	0.73	0.83	0.21	0.60	0.54	0.47

All-by-all trait correlations are listed in Supplementary Table 5

EW, NFSP, NISP, totNSP, NS.NE, SW.EW; mean $h^2 = 0.71$). Individual trait heritabilities were broadly similar between the three trials, with 12 of the 15 traits showing differences in their heritabilities lower than 0.26. The exceptions were 'seed weight-ear weight ratio', 'number of seeds per ear', and 'thousand grain weight', with the former driven by low heritability in UK17 (0.42), and the latter two by relatively low heritability in UK18 (0.46 and 0.61, respectively).

Trait correlations

Correlations between all trait and year combinations are illustrated in Fig. 1, with the underlying values listed in Supplementary Table 5. Correlation within a trait between years was generally high (median = 0.57), with 35 of the 45 between year comparisons (78%) showing a correlation ≥ 0.4 (see also Table 1). As might be expected, the strength of inter-year correlations was itself correlated with trait heritabilities, such that traits with low heritability also tended to have low correlations between years. Overall, traits measured in the two UK trials showed the highest between year correlations (mean = 0.60), followed by trial UK17 versus DE18 and UK18 versus DE18 ($r = 0.55$ and 0.47, respectively), possibly reflecting the contrasting agricultural environments between the sites in the UK and Germany. Particularly strong positive correlations across all three trials were observed for the yield components 'ear length' ($r \geq 0.71$), 'number of fertile spikelets per ear' ($r \geq 0.70$), and 'total number of fertile spikelets per ear' ($r \geq 0.76$), and seed morphometric traits generally showed good positive inter-year correlations, ranging from 0.36 ('seed width', UK18 versus DE18) to 0.73 ('seed length', UK17 versus UK18). In addition to yield component traits, 'flowering time' and 'plant height' both showed high inter-year correlations with values above 0.7 and 0.6, respectively. Notable exceptions to these high correlation trends were the ear traits 'number of infertile spikelets per ear' ($r = 0.14$ to 0.43) and 'seed weight/ear weight ratio' ($r = 0.04$ to 0.16), and to a lesser extent 'ear weight' ($r = 0.22$ to 0.65), for which low correlations were driven by comparisons of the UK versus German trials (Table 1).

Considering correlations between traits within a given trial, the commercially important agronomic trait 'thousand grain weight' was negatively correlated, albeit relatively weakly in some cases, with the 'number of fertile spikelets', 'total number of spikelets' and the 'number of seeds per ear'. This indicated increasing the fertility within the ear comes at the cost of reducing the average size of a given grain amongst the range of grain sizes achieved within a given ear. Notably high positive correlations within each of the three trials were found between 'thousand grain weight' and the grain morphometric traits 'seed area', 'seed width', 'seed length' and 'factor form density' (Fig. 1, Supplementary

Table 5). Furthermore, these traits also showed good positive correlation within all trials with the ear-related traits 'ear weight' and 'ear length'. Conversely, relatively strong negative correlations within each year were observed for the two UK trials between the 'number of infertile spikelets' and the four ear traits ('ear weight', 'ear length', 'number of seed per ear' and 'number of fertile spikelets'), particularly in UK18, which would make sense based on prior assumptions. In the German trial, however, 'number of infertile spikelets per ear' was positively correlated with the 'total number of spikelets per ear' ($r = 0.30$, $p = 4.91^{-8}$), indicating different yield component trade-offs may have taken place in the German growing environment. Indeed, 'number of infertile spikelets' was relatively strongly negatively correlated with 'flowering time' in DE18 ($r = -0.37$, $p = 1.68^{-14}$), while this was not notably the case in the two UK trials. Collectively, these observations indicate different factors between the German trial and those in the UK played a part in driving the 'number of sterile spikelets per ear', as well as the trade-off between the 'number of seeds per ear' and seed morphometric traits. In both UK trials, 'ear weight' was positively correlated with 'plant height' and negatively correlated with 'flowering time', although these correlations were not seen in the German trial.

Genetic analyses

Across all three trials, 187 individual QTL (excluding QTL from the meta-analysis) were detected across the 15 traits investigated: 63 in UK17, 65 in UK18 and 59 in DE18 (Fig. 2; Supplementary Table 6). These QTL were distributed across 19 of the 21 chromosomes of wheat, with fewer located on the wheat D sub-genome (23) than on the A and B sub-genomes (97 and 67, respectively). Within each trial, the average number of QTL per trait was four. The numbers of QTL per trait ranged from two (for 'total number of spikelets') to seven ('seed length') in UK17, from two ('factor form density' and 'weight of seed as a proportion of ear weight') to eight ('seed length') in UK18, and from one ('weight of seed as a proportion of ear weight') to six ('ear weight' and 'seed area') in DE18. Of the 187 individual QTL, 45% were found to be replicated for a given trait in two or more trials, resulting in the identification of 34 genetic loci controlling 13 of the 15 traits investigated (Fig. 2; Supplementary Table 6). The number of replicated QTL per trait ranged from one (for 'number of seeds per ear') to four ('ear length'). No replicated QTL were identified for 'number of infertile spikelets' and 'seed weight-ear weight ratio'.

QTL controlling multiple traits

Of the 34 replicated QTL, 24 were found to resolve into eight genetic loci each controlling ≥ 2 traits, located on
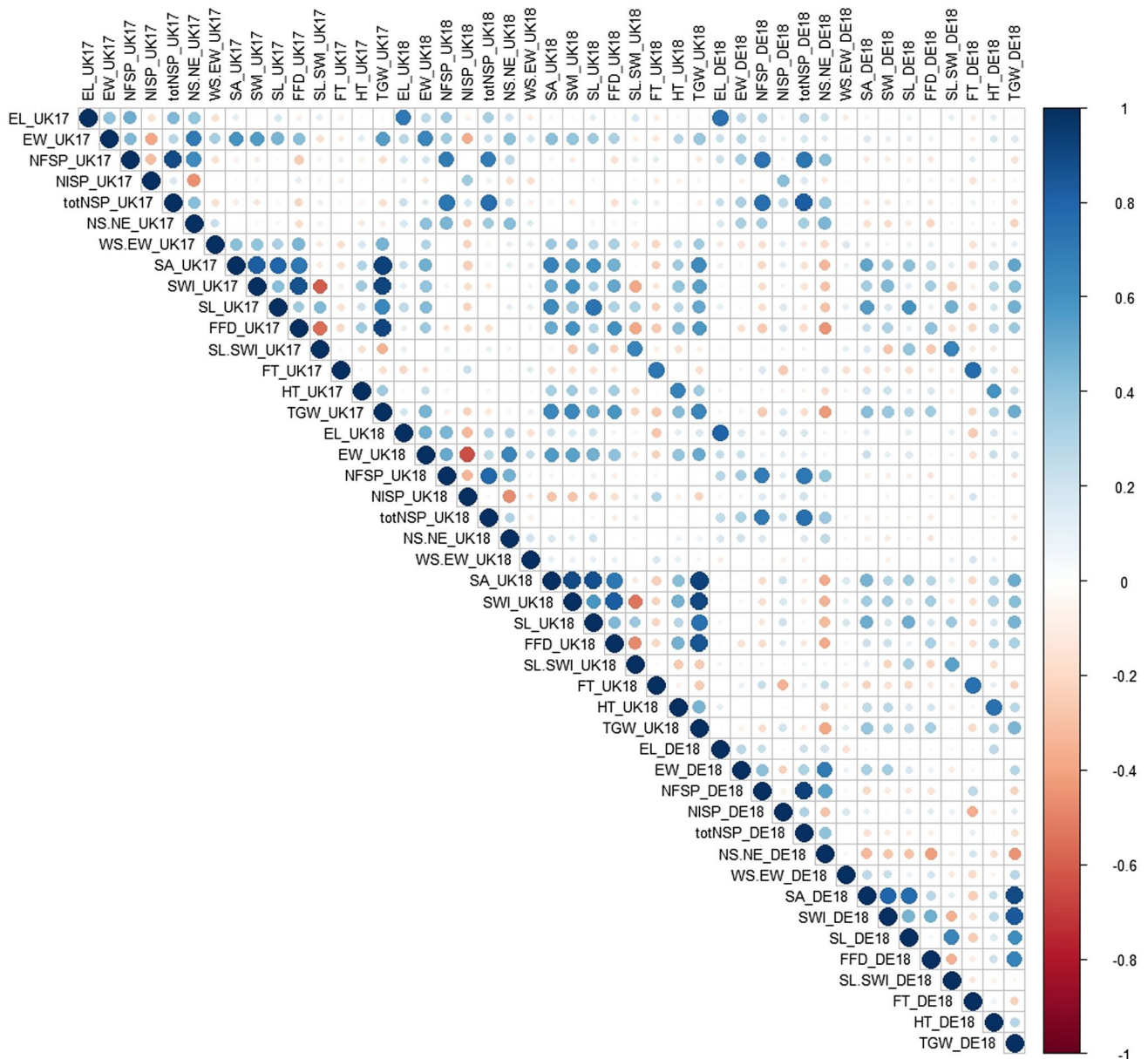
Fig. 1 Correlations between the 15 traits measured in the BMW-pop grown in field trials undertaken in the United Kingdom in 2017 and 2018 (UK17 and UK 18) and in Germany in 2018 (DE18). Trait abbreviations: EL (ear length), EW (ear width), NFSP (number of fertile spikelets per ear), NISP (number of infertile spikelets per ear), totNSP (total number of spikelets per ear), NS.NE (number of seeds per ear), WS.EW (seed weight/ear weight ratio), SA (seed area), SWI (seed width), SL (seed length), FFD (factor form density), SL.SWI (seed length/seed width ratio), FT (flowering time), HT (plant height)

chromosomes 4B, 4D, 5B, 6A, 7A and 7B. Here, these are termed 'multi-trait QTL' (MT-QTL) and named as *QMtqtl.lfl-XX*, where 'XX' indicates chromosome (Fig. 2; Supplementary Table 6). All three of the replicated QTL for 'plant height 'identified in each of our three trials were found to co-locate with MT-QTL (*QMtqtl.lfl-4B.1*, *QMtqtl.lfl-4D.1* and *QMtqtl.lfl-6A.1*). The most significant height QTL identified correspond to the homoeologous *Reduced height* (*Rht*) loci *Rht-B1* and *Rht-D1*, explaining 5.7–9.4 and 14.5–29.9%

of the phenotypic variation for plant height, respectively. These loci have been map-based cloned, and control the semi-dwarfing phenotype of modern wheat that played a central role in the 'Green Revolution' (Peng et al. 1999), and both have previously been shown to segregate in the BMWpop (Stadlmeier et al. 2019) (for BMWpop founder alleles at these loci, see Supplementary Table 7). At *Rht-D1*, replicated QTL for five traits were identified ('plant height', 'seed width', 'seed length–width ratio', 'factor form density

◄**Fig. 2** The genetic map locations of quantitative trait loci (QTL) identified in the BMWpop grown in field trials undertaken in the United Kingdom in 2017 and 2018 (UK17 and UK 18) and in Germany in 2018 (DE18) using interval mapping and composite interval mapping. Multi-trait QTL (MT-QTL) are also included, and are shown in red. 'Strong' QTL, classified as those with -log10*p*-values higher than the trait-specific *p = 0.05* significance thresholds determined by permutation are indicated by solid bars. 'Weak' QTL, classified as those with -$\log_{10}p$-values less than the permutated *p* = 0.05 significance threshold but with -$\log_{10}p > 3$ and which explain ≥ 5% of the phenotypic variation, are indicated by bars with diagonal lines. The BMWpop genetic map is previously published (Stadlmeier et al. 2018)

and 'thousand grain weight'), while two replicated QTL co-segregated at *Rht-B1* ('plant height' and 'ear length'). At each of these two loci, the direction of the predicted allelic haplotype effects for each of the eight founders for 'plant height' were found to be similar to those of the corresponding yield component QTL, indicating pleiotropy rather than genetic linkage was more likely the cause for QTL co-localization (Supplementary Table 6). The third 'plant height' QTL, located on chromosome 6A, controlled 9–12% of the phenotypic variation and co-located with QTL for three grain-related traits ('seed width', 'factor form density' and 'thousand grain weight'). The genetic locus *Rht24* controlling plant height is reported to be located on chromosome 6A (Tian et al. 2017; Würschum et al. 2017; Herter et al. 2018; Scott et al. 2020b). Anchoring the flanking markers of our chromosome 6A 'plant height' QTL to the wheat reference genome (6A: 409–520 Mb, QTL peak at 439 Mb; Supplementary Fig. 2; Supplementary Table 6) found it to span the *Rht24* interval as defined by Würschum et al. (2017) (6A: 400–450 Mb). We therefore subsequently refer to this locus in the BMWpop as *Rht24*. Predicted allelic effects indicated the founders Ambition and Potenzial most consistently conferred a reduction in height, and that the allelic effects for these two founders conferred the strongest reduction at the co-locating QTL for 'seed width', 'factor form density' and 'thousand grain weight', thus indicating height affected these seed traits pleiotropically (Supplementary Table 6). A second MT-QTL was also located on chromosome 6A, *QMtqtl.lfl-6A.2*, consisting of replicated QTL for 'ear weight', 'number of seed per ear', 'seed length' and 'seed area' (Fig. 2, Supplementary Table 6). Predicted founder effects indicated that the founder Format generally had the strongest and most consistent effect on reducing trait expression for all four co-locating traits.

Two MT-QTL were located on chromosome 5B, neither of which were located within the region of low genetic recombination spanning the centromere (Fig. 2, Supplementary Table 6). *QMtqtl.lfl-5B.1* consisted of replicated QTL for 'flowering time' and 'ear length', explaining 5.0–5.9 and 6.5–10.0% of the phenotypic variation, respectively. The trends in the predicted allelic effects at this locus found

founders Firl3565, Potenzial and Bussard to most strongly increase 'ear length'. While alleles at these three founders also resulted in earlier 'flowering time', the effect was not always consistent, and so is not clear whether pleiotropic effects of flowering on ear length occur at this locus. At *QMtqtl.lfl-5B.2*, QTL for two grain morphometric traits colocalized, 'seed length' and 'seed length-seed width ratio', controlling 5.0–7.1% and 7.3–7.9% of the phenotypic variation, respectively. Predicted allelic effects from the founders Event and Format consistently resulted in increased values for both traits, indicating pleiotropy, as might be expected given that 'seed length-seed width ratio' is partially derived from measured values for 'seed length'.

Finally, MT-QTL *QMtqtl.lfl-7A.1* and *QMtqtl.lfl-7B.1* were identified on homoeologous positions on the long arm of chromosomes 7A and 7B, respectively (Fig. 2; Supplementary Table 6). *QMtqtl.lfl-7A.1* encompassed three ear related traits ('number of fertile spikelets', 'total number of spikelets' and 'ear length'), with allelic effects from founders 1, 5, 6 and 7 (Event, Format, Potenzial and Bussard) predicted to increase the numbers of both fertile and infertile spikelets at all sites and years. However, the observed trends for the predicted allelic effects at the 'ear length' QTL differed, with increasing alleles conferred by founders 1, 2, 4, 6 and 7 (Event, BAYP4535, Firl3565, Potenzial and Bussard), indicating the possible presence of a linked QTL controlling 'ear length'. The physical position of our chromosome 7A 'number of fertile spikelets' QTL (peak marker *wsnp_JD_c20555_18262317* at 674.277 Mb) corresponds to the strong effect spikelet number QTL identified in other studies (e.g. 673.779 to 674.277 Mb, Kuzay et al. 2019; 673.780 to 674.300 Mb, Muqaddasi et al. 2019) and was located eight gene models from the wheat gene *WAPO-A1* (*TraesCS7A02G481600*, located at 674.081 Mbp; Supplementary Fig. 3) previously identified as a likely candidate underlying this chromosome 7A QTL (Kuzay et al. 2019; Muqaddasi et al. 2019). We extracted *WAPO-A1* DNA sequences from 16 recently sequenced genome assemblies (Walkowiak et al. 2020), including 1000 bp up- and downstream of the coding regions, identifying 21 DNA variants defining four haplotypes (Supplementary Figs. 3–4; Supplementary Table 8a). Sequencing a 1,719 bp region spanning the *WAPO-A1* coding regions in the eight BMWpop founders (as well as 12 bp and 251 bp up- and down-stream of the start and stop codon, respectively), found the parental lines to possess two of the four haplotypes identified in the 16 lines with sequenced genomes (based on the common 1,719 bp region): *WAPO-A1.hap1* was carried by the four BMWpop founders with low spikelet number alleles at *QMtqtl-7A.1* (Ambition, BAYP4535, Firl3565 and Julius), while *WAPO-A1.hap2* was present in the four founders with high spikelet number alleles (Bussard, Event, Format and Potenzial) (Supplementary Figs. 3–4; Supplementary
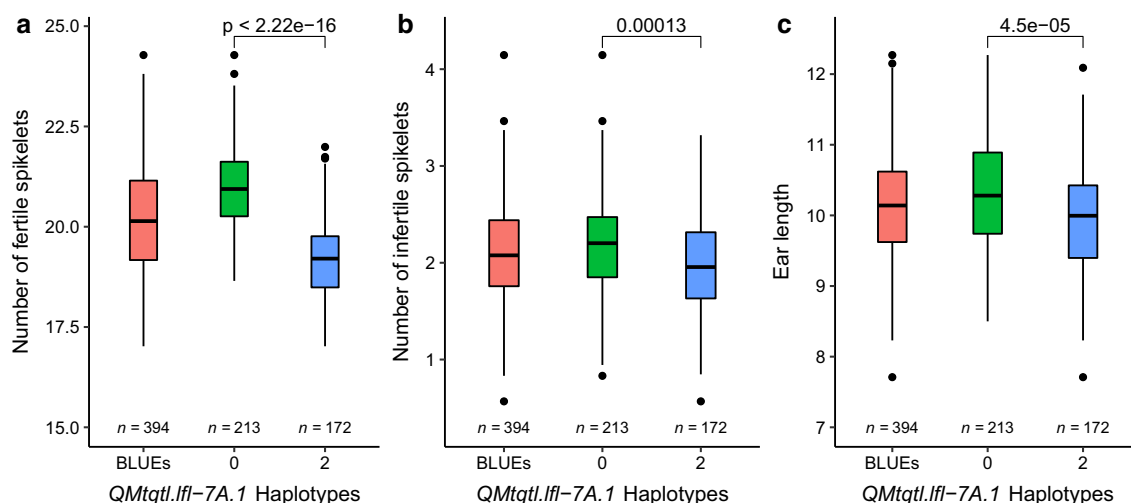
Fig. 3 Analysis of the effects of haplotypes *WAPO-A1.hap1 WAPO-A1.hap2* at the multi-trait QTL *QMtqtl.lfl-7A.1* controlling spikelet number traits on chromosome 7A. Allelic state at SNP *wsnp_JD_c20555_18262317*, located six gene models from *WAPO-A1* in the wheat reference genome (RefSeq v1.0, gene model build RefSeq v1.2; IWGSC 2018), was found to be in phase with the two *WAPO-A1* haplotypes present in the eight BMWpop founders. This SNP was therefore used group the BMWpop recombinant inbred lines (RILs)

for each of the three traits for which QTL co-located at *QMtqtl.lfl-7A.1*: (**a**) 'number of fertile spikelets', (**b**) 'number of infertile spikelets', and (**c**) 'ear length'. For each of these traits, the meta-analysis best linear unbiased estimates (BLUEs) were used for haplotype analysis. The trait values for all 394 BMWpop RILs are shown on the left in red. The significance of the differences between haplotypes was determined by Wilcoxon test ($p < 0.05$)

Table 8a). *WAPO-A1.hap2* included two SNPs that resulted in amino acid substitutions relative to *WAPO-A1.hap1*: a cysteine to phenylalanine at amino acid position 47 (C47/F) located within a highly conserved position in the F-box domain, and an asparagine to aspartate substitution at position 384 (N384/D) at a non-conserved residue toward the C-terminus (Supplementary Figs. 3–4; Supplementary Table 8a), as reported previously by Kuzay et al. (2019). The peak SNP identified at the BMWpop chromosome 7A spikelet number locus (*wsnp_JD_c20555_18262317*) was located seven gene models away from *WAPO-A1*, with the two alleles at SNP *wsnp_JD_c20555_18262317* distinguishing the two *WAPO-A1* haplotypes present in the BMWpop founders. Grouping the BMWpop RILs according to allele call at *wsnp_JD_c20555_18262317* found the RILs predicted to carry *WAPO-A1.hap2* to have a significant increase in the number of fertile spikelets ($P < 2.22e^{-16}$), the number of infertile spikelets ($P = 0.00013$) and ear length ($P = 4.5e^{-05}$), compared to those predicted to carry haplotype *WAPO-A1.hap1* (Fig. 3).

Multi-trait QTL *QMtqtl.lfl-7B.1*, located on chromosome 7B at a homoeologous location to the 7A QTL *QMtqtl.lfl-7A.1*, included co-locating QTL for two related spikelet fertility traits ('number of fertile spikelets' and 'total number of spikelets') that were identified in both UK trials and the meta-analysis (Fig. 4a; Supplementary Table 6). Predicted allelic effects were relatively variable within and between these trait/year combinations, although the effects

were more consistent in the meta-analysis with the founders Firl3565, Potenzial and Julius carrying alleles for higher spikelet number (Supplementary Table 6). Analysis of the IBS results confirmed the IM/CIM result, with significant markers identified for the 7B locus in trials UK17 and UK18 for 'number of fertile spikelets' ($p_{adj} \le 0.0008$) and 'total number of spikelets' ($p_{adj} \le 0.0010$) (Supplementary Table 6). Although genetic analysis using IM/CIM did not identify a significant spikelet number related QTL at the chromosome 7B locus in the DE18 trial, IBS genetic analysis of DE18 identified significant hits for both 'number of fertile spikelets' ($p_{adj} = 0.0060$) and 'total number of spikelets ($p_{adj} = 0.0039$). *WAPO-B1* (*TraesCS7B02G384000*) is located on chromosome 7B at 649.950 Mb, 24 gene models away from the peak markers identified in the IM/CIM meta-analyses for the two spikelet traits (*AX.94949800*, chromosome 7B at 651.340 Mb) (Fig. 4b–c). As the chromosome 7A and 7B spikelet number QTL identified in the BMWpop were located at homoeologous positions, we first investigated the *WAPO-B1* haplotypes present in the 16 hexaploid wheat lines with sequenced genome assemblies used in our analysis of *WAPO-A1*. Extracting *WAPO-B1* genomic sequences, including 1 kb up- and down-stream of the gene, identified regions of missing sequence around a $(CT)_n$ microsatellite upstream of the start codon (Supplementary Fig. 5). We therefore considered the regions starting immediately downstream of this microsatellite in the extracted sequences, identifying 17 polymorphic sites that

defined three haplotypes: *WAPO-B1.hap1, WAPO-B1.hap2* and *WAPO-B1.hap3* (Fig. 4d; Supplementary Fig. 5; Supplementary Table 8b). Sequencing a 1893 bp region spanning *WAPO-B1* in the eight BMWpop founders, and aligning these to the genomic *WAPO-B1* sequences extracted from the 16 lines with genome assemblies, found two of the three *WAPO-B1* haplotypes to be present in the BMWpop. The BMWpop founders carrying the low spikelet number allele at *QMtqtl.lfl-7B.1* all possessed haplotype *WAPO-B1.hap1* (Ambition, Bussard, Event, Format and BAYP4535), while two of the three founders with the high spikelet number allele at the locus carried *WAPO-B1.hap2* (Fig. 4d; Supplementary Table 8b). These two *WAPO-B1* haplotypes were differed to each other by five SNPs: four in exon-1 and one in the 3′ untranslated region (UTR). Of the four exonic SNPs, two did not result in an amino acid change in the predicted protein (G + 594/C and A + 762/G), while the remaining two SNPs resulted in non-synonymous substitutions. The first, G + 140/A, resulted in an arginine to histidine substitution at amino acid residue 47 (R47/H), with alignment of WAPO proteins from 53 plant species showing this substitution to be located in a region of very low sequence alignment close to the N-terminus (Supplementary Fig. 6). The second non-synonymous SNP (T + 517/G) resulted in a serine to alanine substitution at residue 173 (S173/A, BLOSUM62 score = 1, indicating a disruptive amino acid change) within a highly conserved region of the *WAPO-B1* predicted protein at which the S residue was present in 56 of the 53 species investigated (Fig. 4e). We could not amplify *WAPO-B1* in the BMWpop founder Firl3565, despite trying three primer combinations targeting different regions of the gene (Supplementary Table 2). To investigate haplotype effects at *QMtqtl.lfl-7B.1* on spikelet traits in the BMWpop RILs, we constructed a three-SNP haplotype at the QTL using the most significant marker identified in the meta-analysis (*AX.94949800*) and the two flanking SNPs in the genetic map (*Tdurum_contig81911_179* and *Kukri_c12901_706*). The low (designated here *WAPO-B1a*) and high (*WAPO-B1b*) spikelet number alleles identified in the founders were represented by the three-SNP haplotypes '000′ and '222′, respectively. After using the chromosome 7A SNP *wsnp_JD_c20555_18262317* at *QMtqtl.lfl-7A.1* as a cofactor to account for the major effect of this locus on spikelet traits, haplotype analysis of the resulting BLUEs found significant differences between the three-SNP haplotypes at *QMtqtl-lfl-7B.1* for the traits 'number of fertile spikelets' and 'total number of spikelets', but not for 'number of infertile spikelets' (Fig. 4f). Finally, in silico analysis of gene expression found *WAPO-B1* to be expressed in similar tissues to its previously characterized homoeologue *WAPO-A1*, including the shoot apical meristem at different growth stages, with particular high expression at the meristem elongation stage and floret differentiation stage (Fig. 4g; Supplementary

Fig. 7). Unlike *WAPO-A1* however, *WAPO-D1* also showed notably high expression in grain tissues.
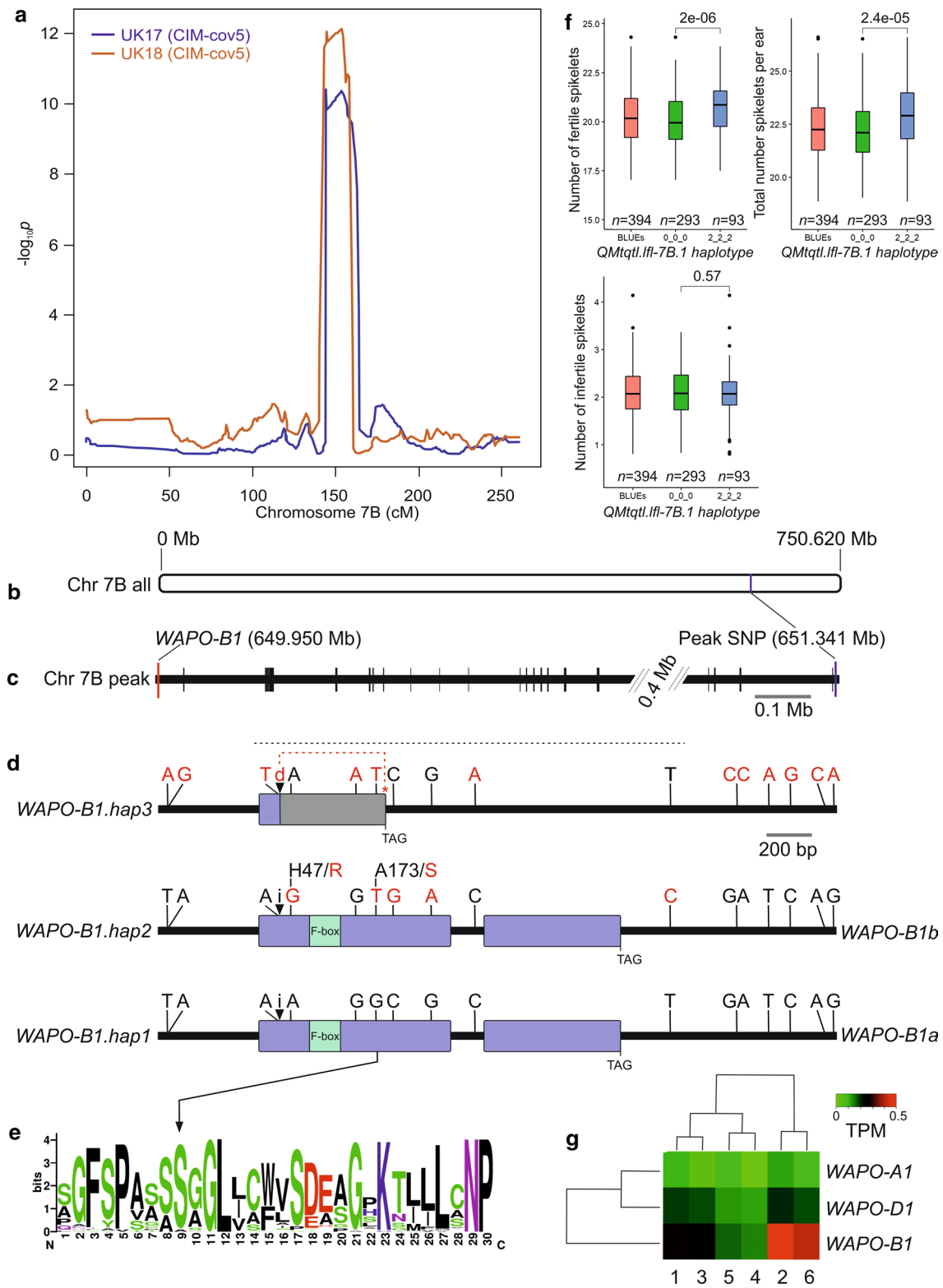
Additional replicated QTL

Nine replicated QTL were identified outside of the regions designated as MT-QTL (Supplementary Table 6). The most significant of these were: (a) the 'seed length/width ratio' QTL *QSlsw.lfl-1B.1* (IM-cov0 -log10p = 5.9–8.0), which explained 7.8–8.5% of the phenotypic variation, and for which the allele from the founder Event was predicted to increase trait expression, and (b) the 'flowering time' QTL *QFt.lfl-3A.1* (IM-cov0 -log10p = 5.1–8.0), explaining 6.8–10.0% of the phenotypic variation and for which alleles from the founders Firl3565 and Format were predicted to delay flowering. Two QTL were identified for 'seed area' at approximately colinear positions spanning the regions of very low genetic recombination that encompass the centromeres on chromosomes 2A (*QSa.lfl-2A.1*) and 2B (*QSa.lfl-2B.1*) (Fig. 2; Supplementary Table 6), explaining around 6% and 5% of the phenotypic variation, respectively. Also located on chromosome 2B was a replicated QTL for 'thousand grain weight' (*QTgw.lfl-2B.1*) explaining 5.9–8.5% of the variation. This QTL also spanned the highly non-recombining region across the chromosome 2B centromere, as did the chromosome 2B 'seed area' QTL (Supplementary Table 6). However, as this region is predicted to contain around half of the physical region of chromosome 2B, we did not classify these QTL as MT-QTL, despite their genetic map locations overlapping to some degree in some year/trial combinations. A 'flowering time' QTL identified in the UK trials, *QFt.lfl-3A.1*, was identified on the long arm of chromosome 3A and explained 6.8–10.0% of the variation. Finally, two replicated QTL were found on chromosome 3B. The first, *QSl.lfl-3B.1*, was located on the short arm and explained 5.2–6.4% of the variation for 'seed length', while the second, *QEw.lfl-3B.1*, was on the long arm and explained ~6% of the variation for 'ear weight'.

Discussion

Trade-offs between key yield component traits and their interaction with plan height

Analysis of the correlations between yield component traits can give insight into the trade-offs that take place between them. Grain number per ear is one of the key components determining final grain yield and is largely set during the early stages of reproductive development (Sreenivasulu and Schnurbusch 2012). Our finding that 'total number of seeds per ear' was negatively associated with seed morphometric traits has been documented by others (e.g. Slafer and

and/or to increased fertility in distal florets within spikelets leading to a higher number of smaller grains per ear. Our observation that the five seed morphometric traits were relatively highly positively correlated across all three trials,

Miralles 1993; Sadras 2007; Bustos et al. 2013). We interpret this as a source limitation on grain yield/size, i.e. in lines where more grains are present per ear, grain filling is limited; possibly due to a lack of sufficient photo-assimilates,

◄Fig. 4 The BMWpop multi-trait QTL *QMtqtl.lfl-7B.1* and details of the underlying candidate gene, *WAPO-B1*. (**a**) QTL on chromosome 7B identified in trials conducted in the United Kingdom in 2017 (UK17) and 2018 (UK18) for the trait 'total number of spikelets', one of the three traits with co-locating QTL at *QMtqtl.lfl-7B.1*. Results of composite interval mapping with five covariates are shown (note, while the significance of the QTL is inflated relative to the results of interval mapping without the inclusion of covariates shown in Supplementary Table 6, the definition of the QTL intervals is improved and so is used here). (**b**) Location on the chromosome 7B physical map (RefSeq v1.0; IWGSC 2018) of the peak SNP marker (*AX.94949800*, in gene model *TraesCS7B02G386400*) identified at QTL for 'total number of spikelets' in trials UK17 and UK18. (**c**) the location *WAPO-B1* (*TraesCS7B02G384000*) on the physical map, 24 high-confidence gene models from the peak SNP identified by QTL mapping. (**d**) *WAPO-B1* haplotypes, based on the DNA variants identified by comparing sequences from the wheat reference genome assembly (IWGSC 2018) with those from 15 additional hexaploid wheat lines with sequenced genomes (Walkowiak et al. 2020). The positions of the 17 DNA variants that define the haplotypes are as detailed in Supplementary Table 8b, with changes in the DNA and amino acid sequences from that of the reference sequence of cv. Chinese Spring indicated in red. The 5 bp insertion (1) / deletion (d) in exon-1 present in *WAPO-B1.hap3* is indicated, and leads to a frame shift (grey) that spans the F-box domain and a subsequent premature stop codon (red asterisk). Sequencing the 1,893 bp region indicated by the dashed horizontal black line in seven of the eight BMWpop founders allowed allocation as *WAPO-B1.hap1* (Ambition, Bussard, Event, Format and BAY4535, all of which carried the low spikelet number allele, designated here *WAPO-B1a*) or *WAPO-B1.hap2* (Julius and Potenzial, allele *WAPO-B1b*); *WAPO-B1* haplotype in the founder Firl3565 is currently undetermined as we were unable to amplify the gene. Of the five SNPs that differentiated the two *WAPO-B1* haplotypes in the seven BMWpop founders, two lead to alterations in the predicted protein in *WAPO-B1.hap2*. The first was an A+140/G SNP that results in a H/47R amino acid substitution in a highly non-conserved region of the protein (alignment shown in Supplementary Fig. 6). The second was a G+517/T SNP resulting in an A173/S amino acid substitution, (**e**) located in a highly conserved region of the protein, based on alignment of 56 proteins from 53 plant species. (**f**) Haplotype analysis at the *QMtqtl.lfl-7B.1* locus in the BMWpop recombinant inbred lines (RILs). Haplotypes were constructed using three SNPs: *Tdurum_contig81911_179*, *Kukri_c12901_706* and *AX.94949800* (the peak marker from the meta-analyses for 'number of fertile spikelets' and 'total number of spikelets'), located at RefSeq v1.0 positions 646,037,877, 646,063,954 and 651,340,580 bp, respectively. Haplotype `000' (representing the combination of the allele calls at each of the three SNPs) tags the low spikelet *WAPO-B1a* allele (Ambition, Bussard, Event, Format, BAYP4535) while haplotype `222' tags the high spikelet *WAPO-B1b* allele (Julius and Potenzial, as well as Firl3565). All BMWpop RILs carried haplotypes `000' or `222', except for two lines carrying haplotype '002' which were excluded for the purposes of the analysis presented here. (**g**) *WAPO-A1, -B1* and *-D1* gene expression in the apical meristem at different stages of development, measured as transcripts per million reads (TPM), sourced from Li et al. (2018b). Apical meristem developmental stages included are numbered in developmental order: 1=vegetative, 2=elongation, 3=single ridge, 4=double ridge, 5=glume primordium, 6=floret differentiation. At stage 4, the meristem has transitioned from the vegetative to reproductive stage

despite having been undertaken in different countries and/or seasons, indicates that genetic improvement specifically targeting these yield components should be robust across broad

geographic areas. Indeed, the heritability of these traits was high across all trials (0.69 to 0.91). Therefore, understanding the potential trade-offs between these seed traits and other yield components could be interesting to investigate further. The finding that four seed morphometric traits ('seed length', 'seed width', 'seed area' and 'factor form density') were positively correlated with plant height across all years and locations indicated that, under our experimental conditions, taller lines supported increased grain size parameters. Indeed, as increased 'plant height' was not generally associated with an increase in spikelet number or number of seed per ear, the effect of this increase in height on grain morphometric traits appears to predominantly promote increased grain size without increasing the overall ear fertility. We found *Rht24* to be the second most important of the three 'plant height' QTL identified after *Rht-D1*, in agreement with previous analysis in the BMWpop (Stadlmeier et al. 2019). Unlike *Rht-B1* and *Rht-D1*, *Rht24* belongs to a group of gibberellic acid (GA) sensitive height altering genetic loci, and the height reducing *Rht24b* allele has been used throughout much of Europe, without notable geographic (and therefore climatic) trend (Würschum et al. 2017). Therefore, it might be assumed that 'tall' alleles at this locus contribute to this increase in grain morphometric traits. In addition to the *Rht24* height locus, QTL for three other traits ('seed width', 'thousand grain weight' and 'factor form density') co-localized within MT-QTL *QMtqtl.lfl-6A.1*. Notably, the wheat gene *TaGW2-A1* (RefSeq v1.1 gene model *TraesCS6A02G189300*, located at 237.74 Mb), orthologous to the rice gene *GRAIN WIDTH 2* controlling grain width and weight (Song et al. 2007), is predicted to lie within the *QMtqtl.lfl-6A.1* genetic interval. In wheat, artificial mutation at *TaGW2* increases grain width (Simmonds et al. 2016; Wang et al. 2018; Zhang et al. 2018a). However, in our population, both *TaGW2-A1* and *Rht24* were located within the region of very low genetic recombination of 0.538 cM that spans the chromosome 6A centromere. As a result, it is difficult for genetic mapping approaches to achieve meaningful genetic resolution for QTL mapping to much of this region. This is exemplified by a recent study that fine-mapped QTL controlling thousand grain weight and grain number per spike to a very tight genetic interval on chromosome 6A of 0.538 cM, with natural variation in the *TaGW2-A1* promotor suggested as responsible for the phenotypic differences observed (Zhai et al. 2018). However, as the 0.538 cM interval also spanned the region of very low genetic recombination spanning the chromosome 6A centromere, this represented a physical interval encompassing around a third of the chromosome (209 Mb, from 236.980 to 445.980 Mb) predicted to contain ~1,700 RefSeq v1.1 gene models, and is comparable to the *QMtqtl.lfl-6A.1* physical interval we identify here. Clearly, map-based approaches using experimental populations will likely find it difficult to

separate the effects of allelic variation at *TaGW2-A1* from those of the many hundreds of genes within this region of low genetic recombination on chromosome 6A—possibly also including *Rht24*, estimated as located between 400–450 Mb at the distal end of the highly non-recombining region (Würschum et al. 2017). Given the well documented pleiotropic effect of *Rht-B1* (Guan et al. 2018) and *Rht-D1* on grain characters (Zhang et al. 2013), and that *Rht24* was the second strongest 'plant height' QTL identified in our population, the balance of evidence suggests that the presence of co-locating grain QTL in this region of chromosome 6A are more likely to be due to pleiotropic effects of *Rht24*. This assumption is supported by previous studies showing that *Rht24* pleiotropically affects thousand grain weight (Tian et al. 2017) as well as by the observations that the only other genetic loci in our population controlling height (*Rht-B1* and *Rht-D1*) also likely have strong pleiotropic effects on seed traits (Fig. 2), and that within each year, the directions of the predicted founder allelic affects at the chromosome 6A locus were similar across all co-locating QTL (Supplementary Table 6). Of course, the recent findings that even the very well characterized wheat genes *Rht-B1* and *Rht-D1* have very closely linked genes controlling thousand grain weight (Xu et al. 2019) and spikelet architecture (Dixon et al. 2018), respectively, mean that further investigation is needed to confirm the extent of any pleiotrophic effect of *Rht24* on yield components.

Confirmation of previously described multi-trait QTL

Identification of previously characterized QTL help confirm the utility of the BMWpop for genetic dissection of yield components. In addition to the identification of the three *Rht* loci discussed above, a major effect QTL controlling spikelet number on the long arm of chromosome 7A was identified, as reported by previous studies (e.g. Quarrie et al. 2006; Würschum et al. 2018; Zhang et al. 2018b; Kuzay et al. 2019; Muqaddasi et al. 2019; Voss-Fels et al. 2019). The recent fine-mapping of this locus by Kuzay et al. (2019), as well as via additional association mapping studies using European varieties (Muqaddasi et al. 2019; Voss-Fels et al. 2019), identified *WAPO-A1* as a strong candidate gene, with a non-synonymous G/T exon-1 SNP present in the high spikelet number allele *WAPO-A1b* leading to a C47/F amino acid substitution in the F-box domain (Kuzay et al. 2019; Muqaddasi et al. 2019). The high spikelet allele at this locus was associated with *WAPO-A1* haplotype 2 (Keyaz et al. 2019). The observation that the peak SNP for our chromosome 7A QTL was seven gene models away from *WAPO-A1*, and that the high spikelet number *WAPO-A1b* allele at this QTL was conferred by founders carrying haplotype *WAPO-A1. hap2*, further demonstrates the utility of the population

for fine-mapping. This chromosome 7A yield component QTL is unusual for wheat, in that it controls a notably large proportion of the phenotypic variance, such that the resulting difference in spikelet number is readily visible by eye. Indeed, analysis across a dataset of predominantly north-western European wheat varieties has previously shown that this locus (based on the location of SNP marker *IAAV5268* at 679.839 Mb) has been under strong breeder selection over the last two decades (Fradgley et al. 2019). We identify effects at this QTL in trials located in two European countries, indicating it is relatively unaffected by prevailing local environmental conditions, as also noted by Voss-Fels et al. (2019). The reason for the strong phenotypic effect of this locus has been suggested by Kuzay et al. (2019) to be because spikelet number is laid down relatively early during spike development (Bonnet 1936), and so is less prone to environmental effects such as heat and lower water availability later in the growing season as the plants mature. Alternatively, it could be due the presence of putative functional mutations in both the promotor (variant 3 in Supplementary Table 8a) and at conserved exonic locations (variants 10 and 15) (Kuzay et al. 2019). Whatever the underlying reason, increasing spikelet number per ear does not on its own necessarily result in overall increases in grain yield, due to pleiotropic effects on other yield components. Indeed, in our MAGIC population we found thousand grain weight and all grain morphometric traits measured to be negatively correlated with increased spikelet number, albeit weakly (Fig. 1). Kuzay et al. (2019) report that the increase in spikelet number afforded by the *WAPO-A1b* allele is associated with an increase in overall grain yield in both an association mapping panel and a bi-parental population, although a decrease in grain weight was also observed in the association panel. It is possible that this contrast in the effect of increased spikelet number on grain weight may be because any panel of wheat varieties investigated has, by definition, been subjected to intense breeder selection for beneficial agronomic traits—principally grain yield. In such genetic backgrounds, the ~15% increase in spikelet number afforded by the *WAPO-A1b* allele may expose the source limitations of the plant, i.e. the ability to supply the developing florets and grain with sufficient resources to maximize grain fill and grain set. In a bi-parental population, such artificial selection to push yield to its limits has not been imposed, and so trade-offs with grain size may not be so apparent. Indeed, we did not find any other grain QTL to co-locate with the chromosome 7A QTL identified in the BMWpop. Having said that, a QTL for increased sterile spikelet number was identified at the chromosome 7A locus, with increased numbers of infertile spikelets associated with increased numbers of fertile spikelets and the *WAPO-A1. hap2* haplotype. Therefore, while BMWpop RILs with

the *WAPO-A1b* allele likely did result in some degree of source limitation, this did not manifest in the identification of QTL for grain morphometric traits at this locus.

Additional multi-trait QTL identified in the BMWpop: chromosome 7B

We identified a multi-trait QTL involved in the control of two spikelet number traits on the long arm of chromosome 7B, at a homoeologous position to the chromosome 7A locus discussed above. However, the chromosome 7B locus explained a much lower amount of the phenotypic variation, ranging between 5 and 7%, and thus more typical of that commonly controlled by wheat yield component QTL. Although studies have identified spikelet number QTL at this location (e.g. Würschum et al. 2018), none to our knowledge have formally determined homoeology to the chromosome 7A locus. The peak SNP for the BMWpop chromosome 7B QTL was located very close to the candidate gene *WAPO-B1*, for which we identified two haplotypes in the BMWpop founders. The recent identification of *WAPO-A1* as underlying the chromosome 7A locus, our finding that alleles conferring higher spikelet number at the chromosome 7B QTL were associated with the *WAPO-B1.hap2* haplotype, and the presence of a non-conservative amino acid substitution in *WAPO-B1.hap2* at a highly conserved region of the protein, supports *WAPO-B1* as a strong candidate gene for the chromosome 7B QTL. That we were unable to amplify *WAPO-B1* in the founder Firl3565 warrants further investigation. However, the Firl3565 allele at the chromosome 7B locus conferred the strongest increase in spikelet number (Supplementary Table 6), and it is not immediately clear how the putative deletion of *WAPO-B1* might lead to this outcome, especially as loss-of-function mutation resulting in premature stop codons in the rice orthologue *APO1* lead to smaller inflorescences (Ikeda et al. 2007). Interestingly, our analysis of *WAPO-B1* genomic sequences in 16 hexaploid wheat lines with sequenced genomes identified *WAPO-B1.hap3* as possessing a 5 bp exonic deletion within the region coding for the F-box domain, leading to a frame shift and a truncated predicted protein. Although not present in the BMWpop, this haplotype was found in four of the 16 lines, including commercial spring wheat varieties from Australia and the International Maize and Wheat Improvement Center (CIMMYT). While further investigation is needed to determine the effect of *WAPO-B1.hap3* on spikelet number, it nevertheless indicates that putative non-functional *WAPO-B1* alleles can be tolerated in commercially viable wheat germplasm, thus not completely excluding the tentative hypothesis that *WAPO-B1* may be deleted in Firl3565. Finally, in silico analysis of *WAPO-B1* gene expression found it to be expressed in the apical meristem during the developmental phases in which final spikelet number is determined, supporting its

role as underlying the spikelet-related QTL identified here. The expression of *WAPO-B1* in the grain tissues, which was particularly strong at the hard dough stage, indicates it may also play additional roles in the regulation of grain development, and warrants further investigation.

Additional multi-trait QTL identified in the BMWpop: chromosome 5B

QTL for 'flowering time' and 'ear length' co-segregated at *QMtqtl.lfl-5B.1*. While flowering time is known to pleiotropically affect many yield component traits, the predicted allelic effects at this locus for these traits did not give a clear picture whether flowering was indirectly affecting ear length. Previous studies in the BMWpop have not identified a flowering time QTL at this location (Stadlmeier et al. 2019). However, although the flowering time QTL identified here by IM/CIM were classified as 'weak', they were identified in all three trials, as well as in the meta-analysis (Fig. 2). The presence of a flowering time QTL at this locus was further confirmed in all three trials by IBS ($p_{adj} \leq 0.0237$) and by IBD in UK17 ($q_{adj} = 0.0317$) genetic analyses (Supplementary Table 6). The higher significance of IBS versus the IBD/CIM mapping methods observed at this QTL may be due to local factors across the locus (such as minor errors in marker order) reducing power of the haplotype-based genetic mapping approaches. Genes controlling flowering time determine the duration of specific developmental phases as the wheat plant matures, and this control is commonly in response to environmental queues (Dixon et al. 2018). Environment is known to affect developing inflorescence length (e.g. Friend 1965), and flowering time affects ear traits such as spikelet number and spikelet morphology (Dixon et al. 2018; Wolde et al. 2019; Worland et al. 1998). While it is feasible that flowering time may have a pleiotropic effect specifically on ear length, we are not aware of any previous reports demonstrating this. Lastly, while the cloned gene *Q* previously found to control wheat ear length is located on chromosome 5B (gene model *TraesCS5A02G473800*; Simons et al. 2006), it is situated > 120 Mb from *QMtqtl.lfl-5B.1* and so does not underlie the locus. For multi-trait QTL *QMtqtl.lfl-5B.2*, the QTL for 'seed length/width ratio' at this locus is presumably due to the co-locating QTL for 'seed length'. While the predicted allelic founder effects for these underlying QTL at *QMtqtl.lfl-5B.2* were not always consistent for all eight founders between years/traits, the allele from the founder Format was notable in that it did consistently confer increased effect for both traits. A previous study using a bi-parental population between Format and another German variety, Pamier, identified a QTL controlling grain size characters at a similar location on chromosome 5B (based on anchoring common SNP marker *IWB39994* to the physical map: 695.660 Mb), with the increasing allele originating

from Format (Mohler et al. 2016). However, while the chromosome 5B locus in the Format x Pamier population controlled thousand grain weight, seed width and seed area, no QTL was found in this population for seed length, making it unclear whether the QTL in these two populations are the same.

Additional multi-trait QTL identified in the BMWpop: chromosome 6A

MT-QTL *QMtqtl.lfl-6A.2* on the long arm of chromosome 6A contained the highest number of constitutive co-locating QTL, including replicated QTL controlling four yield components ('ear weight', 'number of seed per ear', 'seed area' and 'seed length'). All but two of the underlying QTL were identified in the UK trials, indicating the effects at this MT-QTL were predominantly environment specific. It is worth noting, resistance to the fungal diseases yellow rust (in UK18) and septoria tritici blotch (UK17) also co-located at the chromosome 6A locus and explained ∼ 20% and 5% of the variation for disease severity, respectively (Supplementary Table 6). No obvious correlation was seen between the predicted allelic effects at the disease resistance and the yield component QTL. The founder carrying the most resistant yellow rust resistance allele (Potenzial) was also predicted to carry the most susceptible allele for septoria tritici blotch. Furthermore, while the chromosome 6A yellow rust QTL in UK18 co-located with many yield component QTL, the second major yellow rust resistance QTL on chromosome 1A, which also explained ∼ 20% of the variation, did not. While the German trial was sprayed on one occasion with fungicide, this would not have protected the trial completely from disease. Together, these findings indicate that this region of chromosome 6A may contain allelic variation relevant to a number of agronomic traits. Indeed, yellow rust resistance loci have previously been identified at ∼ 600 Mb on chromosome 6A (Muleta et al. 2017; Vazquez et al. 2012), as has a QTL for root architecture (e.g. Alahmad et al. 2019). Interestingly, our meta-analysis identified a flowering time QTL at *QMtqtl.lfl-6A.2*. While previous studies of septoria tritici blotch resistance in the BMWpop have not identified QTL on chromosome 6A, a QTL controlling flowering time has been reported at this location (*QEet.lfl-6A*, 250–252 cM) (Stadlmeier et al. 2019). Flowering plays a role in disease escape for yellow rust infection of the wheat ear (Cromey et al. 1989; Wellings 2003) and septoria tritici blotch (Eriksen et al. 2003; Stadlmeier et al. 2019). Flowering time can also pleiotropically affect yield component traits (reviewed by Cao et al. 2020), as potentially observed here for the flowering time QTL on chromosome 5B. However, the chromosome 6A flowering time QTL explained just 5% of the phenotypic variation, was not identified in the UK17 and DE18 trials,

and the predicted allelic effects were not similar to those seen for the co-segregating yield-component QTL in the UK18 trial. While the balance of evidence suggests possible pleiotropic effects of the chromosome 6A disease resistance QTL on multiple yield component traits at *QMtqtl.lfl-6A.2*, this is by no means clear and warrants further investigation.

Future directions for improved grain yield

Obvious next steps to investigate the QTL identified here would be to exploit the residual heterozygosity present in the BMWpop RILs to develop heterogeneous inbred families (HIFs), so Mendelizing the loci targeted. These HIFs could then be used for downstream investigation, such as detailed physiological characterization of the effects of discreet QTL, differential gene expression and fine-mapping. For multi-trait QTL, HIF materials could be used to further investigate whether co-location of QTL for multiple traits is due to linkage or pleiotropy. For example, by (1) generating recombination events within the target interval to segregate linked loci, or (2) by investigation of HIF materials predicted by to carry different allelic combinations for the target traits. In terms of understanding the genetics of how final grain yield is built across the lifecycle of wheat, approaches such as those employed here that define the genetic architecture controlling multiple constituent yield components can provide entry points into the genes and gene networks controlling yield (Brinton and Uauy 2019). This allows targeted selection for beneficial alleles, the mining of alternative alleles from wider germplasm, and ultimately to the creation of novel alleles at the underlying genes via artificial mutation techniques such as TILLING (Krasileva et al. 2017) and CRISPR/Cas9 gene editing (Shan et al. 2013). There is evidence that in high-yielding agricultural environments, elite cultivars are currently sink limited (Lynch et al. 2017). This indicates that molecular-assisted tracking of genetic loci that increase grains per ear and the manipulation of the genes that underlie them, such as the replicated spikelet number QTL identified here on chromosomes 1B, 3A, 6A, 7A and 7B, could play an important role in helping increase sink strength. A key trade-off in sink traits appears to be between thousand grain weight and spikelet fertility (which itself is composed of numerous traits, including the number of florets that produce grain within a spikelet and the number of spikelets per ear). In north-western Europe, there is evidence for an increasing trend in spikelet fertility over the last > 50 years, while over the same period no such trend is observed for thousand grain weight or spikelet number (Würschum et al. 2018). Therefore, the increase in fertility is likely attributed to the higher occurrence of grain set in more distal florets within the spikelet. At these positions, grain size is much smaller than at positions G1 and G2 (Feng et al. 2018). Carpel size has been reported

to be positively correlated with final grain weight (Hasan et al. 2011; Xi et al. 2015). Whether reduced size in distal florets is due to source limitation during grain filling stage, and/or to reduced carpel size which is established at an earlier stage in plant development is not clear, and would be an interesting target for further exploration in the BMWpop. Either way, increasing grain set and size at G4 could be a clear breeding target to increase both thousand grain weight and grain yield. Exploring ways to achieve this, while simultaneously increasing spikelet number, by for example manipulation of *WAPO* genes on the A, B and D sub-genomes of hexaploid wheat, could therefore provide an approach to push yield potential beyond that currently achieved. Of course, attempts to push the hidden variation in polyploid wheat may have unforeseen consequences: for example, while *TaGW2* double mutants in bread wheat were associated with a greater increase in grain width, length and protein content, triple mutants were found to have wrinkled grain (Zhang et al. 2018a). Nevertheless, the identification of the genes that underlie sink-based yield component QTL, especially those that underlie multi-trait QTL such as those reported here, has promise to unlock further increases in yield potential. Finally, the incorporation of such molecular genetic advances into quantitative genetics methods such as genomic selection (Meuwissen et al. 2001; Jannink et al. 2010) may provide further opportunities for increasing the rate of genetic gain for wheat yield potential. It is likely that genomic selection will be a major source of genetic improvement in wheat breeding in the coming decade (Mackay et al. 2020). Therefore, incorporation of additional molecular genetic data on known genes or genetic loci, such as those identified in our study, may help improve genomic selection for final grain yield. Indeed, methods for the incorporation of such information are now emerging (e.g. Schrag et al. 2018), and exploration of these approaches may help to further increase the rate of genetic gain for wheat yield potential.

Compliance with ethical standards

References

Alahmad S, El Hassouni K, Bassi FM, Dinglasan E, Youssef C, Quarry G, Aksoy A, Mazzucotelli E, Juhász A, Able JA, Christopher J, Voss-Fels KP, Hickey LT (2019) A major root architecture QTL responding to water limitation in durum wheat. Front Plant Sci 10:436

Altschul SF, Gish W, Miller E, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

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

Bonnet OT (1936) The development of the wheat spike. J Agr Res 53:445–451

Brinton J, Simmonds J, Minter F, Leverington-Waite M, Snape J, Uauy C (2017) Increased pericarp cell length underlies a major quantitative trait locus for grain weight in hexaploid wheat. New Phytol 215:1026–1038

Brinton J, Uauy C (2019) A reductionist approach to dissection grain weight and yield in wheat. J Integr Plant Biol 61:337–358

Broman KW, Wu H, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890

Bustos DV, Hasan AK, Reynolds MP, Calderini DF (2013) Combining high grain number and weight through a DH-population to improve grain yield potential of wheat in high-yielding environments. Field Crops Res 145:106–115

Calderini DF, Dreccer MF, Slafer GA (1995) Genetic improvement in wheat yields and associated traits. A re-examination of previous results and the latest trends. Plant Breed 114:108–112

Cao S, Xu D, Hanif M, Xia X, He Z (2020) Genetic architecture underpinning yield component traits in wheat. Theor Appl Genet 133:1811–1823

Cockram J, Chiapparino E, Taylor SA, Stamati K, Donini P, Laurie DA, O'Sullivan DM (2007) Haplotype analysis of vernalization loci

in European barley germplasm reveals novel *VRN-H1* alleles and a predominant winter *VRN-H1/VRN-H2* multi-locus haplotype. Theor Appl Genet 115:993–1001

Cockram J, Mackay I (2018) Genetic mapping populations for conducting high-resolution trait mapping in plants. Adv Biochem Engin/Biotechnol 164:109–138

Cromey MG (1989) Occurrence and effects of stripe rust in wheat spikes in New Zealand. New Zeal J Crop Hort 17:155–158

Cullis BR, Smith A, Coombes N (2006) On the design of early generation variety trials with correlated data. JABES 11:381–393

Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K, Cavanagh C, Swain SM, Boden SA (2018) *TEOSINTE BRANCHED1* regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). Plant Cell 30:563–581

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Fin RD (2019) The Pfam protein families database in 2019. Nucleic Acids Res D1:D427–D432

Eriksen L, Borum F, Jahoor A (2003) Inheritance and localisation of resistance to *Mycosphaerella graminicola* causing septoria tritici blotch and plant height in the wheat (*Triticum aestivum* L.) genome with DNA markers. Theor Appl Genet 107:515–527

FAO (2017) *Online statistical database: Food balance* (Online). FAOSTAT. Available: http://www.fao.org/faostat/en/ Accessed 28 October 2020.

Faralli M, Cockram J, Ober E, Wall S, Galle A, Van Rie J, Raines C, Lawson T (2019) Genotypic, developmental and environmental effects on the rapidity of g_s in wheat: impacts on carbon gain and water-use efficiency. Front Plant Sci 10:492

Feng F, Han Y, Wang S, Yin S, Peng Z, Zhou M, Gao W, Wen X, Qin X, Siddique KHM (2018) The effect of grain position on genetic improvement of grain number and thousand grain weight in winter wheat in North China. Front Plant Sci 9:129

Fradgley N, Gardner KA, Cockram J, Elderfield J, Hickey JM, Howell P, Jackson R, Mackay I (2019) A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. PLoS Biol 17:e3000071

Friend DJC (1965) Ear length and spikelet number of wheat grown at different temperatures and light intensities. Can J Bot 43:345–353

Garcia M, Eckermann P, Haefele S, Satija S, Sznajder B, Timmins A, Baumann U, Wolters P, Mather DE, Fleury D (2019) Genome-wide association mapping of grain yield in a diverse collection of spring wheat (*Triticum aestivum* L.e) evaluated in southern Australia. PLoS ONE 14:e0211730

Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape variation in wheat. Plant Cell 22:1046–1056

Golan G, Ayalon I, Perry A, Zimran G, Ade-Ajayi T, Mosquna A, Distelfeld A, Peleg Z (2019) *GNI-A1* mediates trade-off between grain number and grain weight in tetraploid wheat. Theor Appl Genet 132:2353–2365

Griffiths S, Simmonds J, Leverington M et al (2009a) Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. Theor Appl Genet 119:383–395

Griffiths S, Simmonds J, Leverington M, Wang Y, Fish L, Sayers L, Alibert L, Orford S, Wingen L, Snape J (2009b) Meta-QTL analysis of the genetic control of crop height in European winter wheat. Mol Breed 29:159–171

Guan PF, Lu LH, Jia LJ, Kabir MR, Zhang JB, Lan TY, Zhao Y, Xin MM, Hu ZR, Yao YY, Ni ZF, Sun QX, Peng HR (2018) Global QTL analysis identifies genomic regions on chromosomes 4A and 4B harboring stable loci for yield-related traits across different environments in wheat (*Triticum aestivum* L.). Front Plant Sci 9:529

Harrell FE with contributions from Charles Dupont and many others (2019) Hmisc: Harrell Miscellaneous. R package version 4.2–0. https://CRAN.R-project.org/package=Hmisc. Accessed 26 October 2020.

Hasan AK, Herrera J, Lizana C, Calderini DF (2011) Carpel weight, grain length and stabilized grain water content are physiological drivers of grain weight determination of wheat. Field Crops Res 123:241–247

Herter CP, Ebmeyer E, Kollers S, Korzun V, Leiser WL, Würschum T, Miedaner T (2018) *Rht24* reduces height in the winter wheat population 'Slitar x Bussard' without adverse effects on Fusarium head blight infection. Theor Appl Genet 131:1263–1272

HoweContreras-Moreira KLB, De Silva N, Maslen G, Akanni W et al (2019) Ensembl Genomes 2020—enabling non-vertibrate genomic research. Nucleic Acids Res 48:D689–D695

Huang BE, George AW (2011) R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. Bioinformatics 27:727–729

Hughes N, Oliveira HR, Fradgley N, Corke FMK, Cockram J, Doonan JH, Nibau C (2019) µCT trait analysis reveals morphometric differences between domesticated temperate small grain cereals and their wild relatives. Plant J 99:98–111

Ikeda K, Ito M, Nagasawa N, Kyozuka J, Nagato Y (2007) Rice *ABERRANT PANICLE ORGANIZATION 1*, encoding an F-box protein, regulates meristem fate. Plant J 51:1030–1040

The International Wheat Genome Sequencing Consortium (IWGSC), Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, Pozniak CJ, Stein N, Choulet F, Distelfeld A, Eversole K, Poland J, Rogers J, Ronen G, Sharpe AG, Pozniak C, Ronen G, Stein N, Barad O, Baruch K, et al. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361: eaar7191

Jannink JL, Lorenz A, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 9:166–177

Kassambara A (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.2.5. https://CRAN.R-project.org/package=ggpubr.

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H et al (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci U S A 104:1424–1429

Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramírez-González RH, Wang X, Borrill P, Fosker C, Ayling S, Phillips AL, Uauy C, Dubcovsky J (2017) Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci U S A 114:E913–E921

Kuzay S, Xu Y, Zhang J et al (2019) Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. Theor Appl Genet 132:2689–2705

Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X (2018a) A genome-wide view of transcriptome dynamics during early spike development in bread wheat. Scientific Reports 8:15338

Li N, Xu R, Duan P, Li Y (2018b) Control of grain size in rice. Plant Reprod 31:237–251

Lin M, Stadlmeier M, Mohler V, Tan K-C, Ficke A, Cockram J, Lillemo M (2020) Identification and cross-validation of genetic loci conferring resistance to Septoria nodorum blotch using a German multi-founder winter wheat population. Theor Appl Genet 134:125–142

Lynch JP, Doyle D, McAuley S, McHardy F, Danneels Q, Black LC, White EM, Spink J (2017) The impact of variation in grain number and individual grain weight on winter wheat yield in the high yield potential environment of Ireland. Eur J Agron 87:40–49

Ma J, Tu Y, Zhu J, Luo W, Liu H, Li C et al (2020) Flag leaf size and posture of bread wheat: genetic dissection, QTL validation and their relationships with yield-related traits. Theor Appl Genet 133:297–315

Mackay IJ, Cockram J, Howell P, Powell W (2021) Understanding the classics: the unifying concepts of transgressive segregation, inbreeding depression and heterosis and their central relevance for crop breeding. Plant Biotech J 19:26–34

Madeira F, Park YM, Lee J, Buso N, Madhusoodanan N, Basutkar P, Tivey ARN, Poter SC, Finn RD, Lopez R (2019) The EMBL_EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641

McIntosh RA, Yamazaki Y, Dubcovsky J, Rogers J, Morris C, et al. (2013) Catalogue of gene symbols for wheat. In: 12th International Wheat Genetic Symposium. Yokohama, Japan, September

de Mendiburu F (2009) Una herramienta de analisis estadistico para la investigacion agricola. Universidad Nacional de Ingenieria (UNI-PERU), Tesis

Meuwissen T, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Mohler V, Albrecht T, Castell A, Diethelm M, Schweizer G, Hartl L (2016) Considering causal genes in the genetic dissection of kernel traits in common wheat. J Appl Genetics 57:467–476

Mosleth EF, Lillehammer M, Pellny TK, Wood AJ, Riche AB, Hussain A, Griffiths S, Hawkesford MJ, Shewry PR (2020) Genetic variation and heritability of grain protein deviation in European wheat genotypes. Field Crops Res 255:107896

Muleta KT, Bulli P, Rynearson S, Chen X, Pumphrey M (2017) Loci associated with resistance to stripe rust (Puccinia striiformis f sp tritici) in a core collection of spring wheat (*Triticum aestivum*). PLoS ONE 12:e0179087

Muqaddasi QH, Brassac J, Koppolu R et al (2019) TaAPO-A1, an ortholog of rice ABERRANT PANICLE ORGANIZATION 1, is associated with total spikelet number per spike in elite European hexaploid winter wheat (*Triticum aestivum* L.) varieties. Sci Rep 9:13853

Murai K, Takumi S, Koga H, Ogihara Y (2002) Pistillody, homoeotic transformation of stamens into pistil-like structures, caused by nuclear-cytoplasm interaction in wheat. Plant J 29:169–181

Nicholas KB, Nicholas HBJ, Deerfield DWI (1997) GeneDoc: analysis and visualization of genetic variation. Embnew News 4:14

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM et al (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261

Piepho H-P, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888

Quarrie SA, Quarie SP, Radosevic R, Rancic D, Kaminska A, Barnes JD et al (2006) Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. J Exp Bot 57:2627–2637

R Core Team (2014) A Language and Environment for Statistical Computing. R Found. Stat. Comput

Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J et al (2018) The transcriptional landscape of polyploid wheat. Science 361:eaar6089

RStudio Team (2020) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Rustgi S, Shafwat MN, Jumar N, Baenziger PS, Ali ML, Dweikat I, Campbell BT, Gill KS (2013) Genetic dissection of yield and its component traits using high-density composite map of wheat chromosome 3A: bridging gaps between QTLs and underlying genes. PLoS ONE 8:e70526

Sadras VO (2007) Evolutionary aspects of the trade-off between seed size and number in crops. Field Crops Res 100:125–138

Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A et al (2019) Unleashing floret fertility in wheat through the mutation of a homeobox gene. Proc Natl Acad Sci U S A 116:5182–5187

Samach A, Klenz JE, Kohami SE, Risseeuw E, Haughn GW, Crosby WL (1999) The *UNUSUAL FLORAL ORGANS* gene of *Arabidopsis thaliana* is an F-box protein required for normal patterning and growth in the floral meristem. Plant J 20:433–445

Schrag TA, Westhues M, Schipprack W, Seifert F, Thiemann A, Scholten S, Melchinger AE (2018) Beyond genomic prediction: combining different types of omics data can improve prediction of hybrid performance in maize. Genetics 208:1373–1385

Scott MF, Ladejobi O, Amer S, Bentley AR, Biernaskie J et al (2020a) Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with pre-breeding. Heredity 125:396–416

Scott MF, Fradgley N, Bentley AR, Brabbs T, Corke F, Gardner KA, Horsnell R, Howell P, Ladejobi O, Mackay IJ, Mott R, Cockram J (2020b) Limited haplotype diversity underlies polygenic trait architecture across 70 years of wheat breeding. bioRxiv 2020.09.15.296533

Simmonds NW (1995) The relationship between yield and protein in cereal grain. J Sci Food Agric 67:309–315

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z et al (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:686

Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, Blanco A, Dubcovsky J, Uauy C (2016) A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor Appl Genet 129:1099–1112

Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai YS, Gill BS, Faris JD (2006) Molecular characterization of the major wheat domestication gene *Q*. Genetics 172:547–555

Slafer GA, Miralles DJ (1993) Fruiting efficiency in three bread wheat (*Tritkum aestivum*) cultivars released at different eras. Number of grains per spike and grain weight. J Agronomy Crop Sci 170:251–260

Song X-J, Huang W, Shi M, Shu M-Z, Lin H-X (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature Genet 39:623–630

Sreenivasulu N, Schnurbusch T (2012) A genetic playground for enhancing grain number in cereals. Trends Plant Sci 17:91–101

Stadlmeier M, Hartl L, Mohler V (2018) Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. Front Plant Sci 9:1825

Stadlmeier M, NistrupJørgensen L, FejerJustesen A, Corsi B, Cockram J, Hartl L, Mohler V (2019) Genetic dissection of resistance to the three fungal plant pathogens *B. graminis*, *Z. tritici*, and *P. tritci-repentis* in the background of a multiparental winter wheat population. G3 Genes Genomes Genet 9:1745–1757

Tian Z, Wen W, Xie L, Fu L, Xu D, Fu C, Wang D, Chen X, Xia X, Chen Q, He Z, Cao S (2017) Molecular mapping of Reduced Plant Height gene *Rht24* in bread wheat. Front Plant Sci 8:1379

Vazquez MD, Peterson CJ, Riera-Lizarazu O, Chen XM, Heesacker A, Ammar K, Crossa J, Mundt CC (2012) Genetic analysis of adult plant quantitative resistance to stripe rust in wheat cultivar 'Stephens' in multi-environment trials. Theor Appl Genet 124:1–11

Voss-Fels KP, Keeble-Gagnère G, Hickey LT, Tibbits J, Hayden M, Pasam RK, Friedt W, Snowdon RJ, Appels R, Wittkop B (2019) High-resolution mapping of rachis nodes per rachis, a critical determinant of grain yield components in wheat. Theor Appl Genet 132:2707–2719

Walkowiak S, Gao L, Monat C, Haberer G, Kassa MT, Brinton J et al (2020) Multiple wheat genomes reveal global variation in modern breeding. Nature 588:277–283

Wang W, Simmonds J, Pan Q, Davidson D, He F, Battal A, Akhunova A, Trick HN, Uauy C, Akhunov E (2018) Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of *TaGW2* homoeologues to grain size and weight in wheat. Theor Appl Genet 131:2463–2475

Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, et al. (2012) Package 'gplots'. http://cran.r-project.org.

Wei T, Simko V (2017) R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot

Wellings CR (2003) Stripe Rust Head Infection. Cereal Rust Report 1:1–2

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

Wilkinson MD, Haughn GW (1995) *UNUSUAL FLORAL ORGANS* controls meristem identity and organ primordia fate in Arabidopsis. Plant Cell 7:1485–1499

Wolde GM, Trautewig C, Mascher M, Schnurbusch T (2019) Genetic insights into morphometric inflorescence traits of wheat. Theor Appl Genet 132:1661–1676

Worland A, Börner A, Korzun V, Li W, Petrovic S, Sayers E (1998) The influence of photoperiod genes on the adaptability of European winter wheats. Euphytica 100:385–394

Würschum T, Langer SM, Longin CFH, Tucker MR, Leiser WL (2017) A modern Green Revolution gene for reduced height in wheat. Plant J 92:892–903

Würschum T, Leiser W, Langer SM, Longin CFH, Tucker MR, Longin CFH (2018) Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. Theor Appl Genet 131:2071–2084

Xie Q, Mayes S, Sparkes DL (2015) Carpel size, grain filling, and morphology determine individual grain weight in wheat. J Exp Bot 66:6715–6730

Xu D, Wen W, Fu L, Li F, Li J, Xie L, Xia X, Ni Z, He Z, Cao S (2019) Genetic dissection of a major QTL for kernel weight spanning the *Rht-B1* locus in bread wheat. Theor Appl Genet 132:3191–3200

Xue S, Xu F, Li G et al (2013) Fine mapping *TaFLW1*, a major QTL controlling flag leaf width in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 126:1941–1949

Zhai H, Feng Z, Du Xm SY, Liu X, Qi Z, Song L, Li J, Li L, Peng H, Hu Z, Yao Y, Xin M, Xiao S, Sun Q, Ni Z (2018) A novel allele of *TaGW2-A1* is located in a finely mapped QTL that increases grain weight but decreases grain number in wheat (*Triticum aestivum* L.). Theor Appl Genet 131:539–553

Zhang J, Dell B, Biddulph B, Drake-Brockman F, Walker E, Khan N, Wong D, Hayden M, Appels R (2013) Wild-type alleles of *Rht-B1* and *Rht-D1* as independent determinants of thousand grain weight and kernel number per spike in wheat. Mol Breed 32:771–783

Zhang Y, Li D, Zhang D, Zhao X, Cao X, Dong L, Liu J, Chen K, Zhang H, Gao C, Wang D (2018) Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits. Plant J 94:857–866

Zhang JL, Gizaw SA, Bossolini E, Hegarty J, Howell T, Carter AH, Akhunov E, Dubcovsky J (2018) Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. Theor Appl Genet 131:1741–1759