

RESEARCH

Open Access



Ectopic expression of *TaBG1* increases seed size and alters nutritional characteristics of the grain in wheat but does not lead to increased yields

Matthew J. Milner^{*}, Sarah Bowden, Melanie Craze and Emma J. Wallington

Abstract

Background: Grain size is thought to be a major component of yield in many plant species. Here we set out to understand if knowledge from other cereals such as rice could translate to increased yield gains in wheat and lead to increased nitrogen use efficiency. Previous findings that the overexpression of *OsBG1* in rice increased yields while increasing seed size suggest translating gains from rice to other cereals may help to increase yields.

Results: The orthologous genes of *OsBG1* were identified in wheat. One homoeologous wheat gene was cloned and overexpressed in wheat to understand its role in controlling seed size. Potential alteration in the nutritional profile of the grains were also analyzed in wheat overexpressing *TaBG1*. It was found that increased *TaBG1-A* expression could indeed lead to larger seed size but was linked to a reduction in seed number per plant leading to no significant overall increase in yield. Other important components of yield such as biomass or tillering did not change significantly with increased *TaBG1-A* expression. The nutritional profile of the grain was altered, with a significant decrease in the Zn levels in the grain associated with increased seed size, but Fe and Mn concentrations were unchanged. Protein content of the wheat grain also fell under moderate N fertilization levels but not under deficient or adequate levels of N.

Conclusions: *TaBG1* does control seed size in wheat but increasing the seed size *per se* does not increase yield and may come at the cost of lower concentrations of essential elements as well as potentially lower protein content. Nevertheless, *TaBG1* could be a useful target for further breeding efforts in combination with other genes for increased biomass.

Keywords: Seed size, Big grain 1, Nitrogen, Zinc, Iron, Seed number, Auxin

Background

Improving crop yield while maintaining levels of inputs such as nitrogen (N) fertilizer, is necessary to ensure food security for the projected increase in population in years to come. Wheat (*Triticum aestivum* L.) is one of three main staple crops worldwide grown on every continent,

except Antarctica, and is planted on over 200 million hectares annually [1]. Unfortunately wheat yields in many parts of the world have plateaued since the mid-1990 s [2]. This yield plateau can be attributed partly to inadequate agronomic practices, however with the release of an increasingly better reference wheat genome, gains from reverse genetics shown in other model species or crops can now be translated to wheat [3–6]. More recent progress in yield enhancement in wheat has been attributed to the increase in spikelet fertility, or the number

*Correspondence: matthew.milner@niab.com
NIAB, 93 Lawrence Weaver Road, Cambridge CB3 0LE, UK



of kernels per spikelet. This led to the number of seeds per spike and spikelet fertility also being correlated with increased yield [7]. Unfortunately, in wheat both seed number and spikelet fertility showed a negative correlation with thousand-grain weight [7]. Combining both spikelet fertility and seed size may help to drive further yield enhancements. In wheat, since single grain weight varies depending on growth conditions, yield is affected by both the grain number and size [8, 9]. Thus, the targets for a high yield may be more complicated in wheat than in rice where grain number is the major limitation to yield [10, 11]. Thus a greater understanding of components of grain size can lead to improvement in the genetic stock of the cultivars grown, leading to further yield improvements for farmers.

Grain size and weight, which is determined by its three-dimensional structure (length, width, and thickness) is a crucial component that affect yield, quality and nutrition [12]. In the context of crop domestication and artificial breeding, grain size and weight have always been important agronomic traits for selection. One of the major components of increasing grain yield in all crop species is increased seed size. In several plant species an increase in seed size led to increased yields either per plant or in the field. Advancements in our understanding of the genes and or regions of the genome involved in the control of grain size continues to increase. There are several genes which have been identified as playing a role in the modification of seed size which first were identified in rice and later translated to wheat as genetic resources became available. This includes modification of pathways involved in proteasomal degradation (GW2 and GW5/qSW5), phytohormone signaling (OsGS6, TGW6, OsBG1, and OsCKX2), and G protein-mediated signal transduction (GS3 and RGB1); which are thought to work by regulating cell division and/or cell expansion in specific grain tissues [13].

Several wheat genes have also now been isolated and suggested to be associated with thousand grain weight (TGW). These include TaCwi [14, 15], TaTGW6 [16, 17], TaGW2-A1 [18–20], TaGS1a [21], TaGS5-3 A [22], TaGASR7-A1 [23], TaPSTOL [24], TaFlo2 [25], TaCYP78A3 [26], and TaCYP78A5 [27], da1 [28]. Many of these genes have trade-offs with other yield components which ultimately limit their ability to increase yields alone. Increasing our knowledge of genes and pathways involved in seed size and shape are important to understand their role in yield prior to incorporation into breeding programs and the development of new higher yielding varieties.

Other considerations for the growth of a sustainable crop is the amount of inputs required and their efficient utilization by the crop. Nitrogen is one of the most

yield-limiting nutrients in crop production, and a proper understanding of its role in yield components is essential for improving grain yield. In both rice and wheat, grain number is linearly correlated with total plant N content. It has been suggested that most of the variation in both grain number and yield is caused by differences in resource accumulation, particularly nitrogen, by the crop [29]. Application of nitrogen fertilizers prior to the panicle initiation stage in rice is known to affect inflorescence development and increase flower numbers per panicle which is a major driver of increasing yield [10]. There is evidence that rice and wheat may behave differently in their response to N fertilizer, including how they partition N in leaves of the canopy and its effect on yield [30]. This understanding has led to studies where modification of genes involved in increasing seed size does not always lead to an increase in yield in wheat [18, 19]. Other limitations in the translation of genes from rice to wheat is that the architecture of the rice inflorescence and wheat ear are quite different. Some of the genes which have been shown to alter grain size and shape, such as GS3, which has a multi-domain structure, can change the effect it has on seed size and plant height, depending on where the mutations are located [31–33]. Other genes may also alter many potential yield components in addition to seed size, or be specific to the way the crop matures and if the crop has an indeterminate growth habit [33, 34]. Few characterized genes which increase seed size, positively impact on overall yield without pleiotropic effects.

Of major concern is the potential for a reduction in nutritional quality of the grain with increased grain size. As most of the micronutrients are located in the aleurone layer, a change in the endosperm to aleurone layer ratio can reduce the nutritional quality of the consumed food product. This is of serious concern in many countries including the UK, where over half of girls aged 11 to 18 are considered anemic [35]. Thus a reduction in the amount of Fe in the grain could reduce intakes further in an already Fe deficient population. Indeed over the past sixty years the levels of Fe, Zn, Cu and Mg in the grain have decreased significantly as yield has doubled, coinciding with the introduction of semi-dwarf, high-yielding varieties of wheat [36].

Wheat is also a large component of protein in the human diet, with cereals and cereal products providing more than 20% of the protein in the diet in the UK for children over 1.5 years and adults [37]. It is also known that N fertilization levels can influence protein content with typical ranges of protein content ranging from 10 to 15% but can be as high as 22% in some wheat collections [38, 39]. If less than optimum N application is applied to a crop, a lower protein content or reduced yields would

likely occur. This reduction in protein or yield can have both a nutritional and economic costs affecting overall human health.

To understand the role increased seed size could have on yield and nutritional traits we studied the effects of increased *TaBG1* expression in wheat and its role to shape seed size, yield components, and nutritional qualities.

Methods

Gene identification

The *OsBG1* coding sequence from NCBI (Q10R09.1) also known as Os03g0175800 was used as a query for Big Grain genes in Ensembl plants. (<http://plants.ensembl.org/index.html>). Protein domain identification was found using the biomaart section of Ensembl plants v42 to search both the rice and wheat genomes for genes using the Interpro ID IPR039621. All DNA and Proteins sequences are available at Ensembl plants. Expression profiles of the Big Grain like genes was downloaded from the Wheat Expression Browser, (<http://www.wheat-expression.com>) [40]. Phylogenetic tree was constructed using the amino acid sequences of the Big Grain genes in MEGA X [41]. Sequences were aligned using the MUSCLE algorithm and the tree was constructed using 500 iterations using the Maximum Likelihood Tree option. Heat map of expression values was constructed using heatmapper (<http://www.heatmapper.ca/>) [42].

Wheat transformation

Wheat variety Fielder (USDA, available from CIM-MYT, under the name BW 35215) plants were grown in controlled environment chambers (Conviron) at 20°C day/15°C night with a 16 h day photoperiod (approximately 400 $\mu\text{E m}^{-2} \text{s}^{-1}$). Immature seeds were harvested for transformation experiments at 14–20 days post-anthesis (dpa). Isolated immature wheat embryos were co-cultivated with *Agrobacterium tumefaciens* for 2 days in the dark [43]. Subsequent removal of the embryonic axis and tissue culture was performed as described by [44]. Individual plantlets were hardened off following transfer to Jiffy-7 pellets (Jiffy), potted up into 9 cm plant pots containing M2 and grown on to maturity and seed harvest in controlled environment chambers, as above.

Plasmid construction for genetic modification

TaBG1-A was synthesized from the public sequence available at the time for the wheat cultivar Chinese Spring with attL1 and attL2 sites for direct recombination into binary gateway vectors. *TaBG1-A* was then recombined into the binary vector pSc4ActR1R2 using a Gateway LR Clonase II Kit (ThermoFisher) to create pMM14. *TaBG1-A* was expressed in wheat cv Fielder from the rice *Actin*

promoter and transcripts terminated by the *Agrobacterium tumefaciens* nopaline synthase terminator (tNOS).

pMM14 was verified by restriction digest and sequencing before being electro-transformed into *A. tumefaciens* strain EHA105. Plasmids were re-isolated from *Agrobacterium* cultures and verified by restriction digest prior to use in wheat experiments [45].

DNA analysis of transformed wheat plants

Plantlets which regenerated under G418 selection in tissue culture were transferred to Jiffy-7 pellets and validated using an *nptII* copy number assay relative to a single copy wheat gene amplicon, GaMyb, normalized to a known single copy wheat line. Primers and Taqman probes were used at a concentration of 10 μM in a 10 μl multiplex reaction using Absolute Blue qPCR ROX mix (ThermoFisher) with the standard run conditions for the ABI 7900 HT. The relative quantification, $\Delta\Delta\text{CT}$, values were calculated to determine *nptII* copy number in the T_0 and subsequent generations [46]. Homozygous transgenic lines were identified on the basis of *nptII* copy number and segregation analysis. WT Fielder plants were null segregants.

Expression of TaBG1 in wheat

Total RNA was isolated from wheat shoots using a RNeasy Kit (Qiagen). Following treatment with DNaseI (ThermoFisher), cDNA synthesis was conducted on 500 ng of total RNA using Omniscript RT Kit (Qiagen). The cDNA was diluted 1:2 with water and 0.5 μl was used as template in each RT-PCR reaction. Transcript levels were quantified using SYBR Green JumpStartTaq ReadyMix (SIGMA) with the standard run conditions for the ABI 7900 HT. Three technical replicates were performed on each of the three biological replicates. Two reference genes *TaUbiquitin* and *TaEF1 α* were used for the normalization using the $\Delta\Delta\text{CT}$. The sequence of primers used in Q-PCR assays are *TaUbiquitin* F (5'-CCTTCACTTGGTTCTCCGTCT-3'), *TaUbiquitin* R (5'-AACGAC CAGGACGACAGACACA-3'), *TaEF1 α* F (5'-TGGTGT CATCAAGCCTGGTATGGT-3'), *TaEF1 α* R (5'-ACT CATGGTGCATCTCAACGGACT-3'), *TaBG1* F (5'-GCT GCTGGACGCGATATAC-3') and *TaBG1* R (5'-CTG CTTCTTGGCCTTCTTCT-3').

Growth conditions

Transgenic lines and corresponding null segregants were grown on TS5 low fertility soil to control total nitrogen with a starting nitrogen level of 0.1 mg/l (Bourne Amenity, Kent, UK). Ammonium nitrate was then added to each pot to reach a final concentration in the pots equivalent to field fertilizer application of 70, 140 or 210 kg N / ha. Plants were grown in a climate controlled glasshouse

with 10,000 lx sodium supplemental light for a 16 h day and 20°C/15°C day night temperatures. At least 21 plants per line were grown per treatment in a randomized design.

Grain cross sections

The wheat grains were soaked in distilled water and subjected to a vacuum system for 10 min. The seeds were then fixed overnight with 3% EM grade glutaraldehyde in 0.1 M Cacodylate buffer and oscillated overnight at room temperature. Seeds were washed twice in 0.1 M cacodylate buffer for 10 min for each wash and then stored in 0.1 M Cacodylate buffer prior to post tissue-fixation with 1% Osmium Tetroxide in 0.1 M Cacodylate Buffer for two days at 4 °C.

Seeds were then washed twice in distilled water for 10 min each time and dehydrated with 50%, 70%, 90% (2 × 15 min for each) and 100% ethanol (3 × 30 min) respectively. To further remove water from the samples, propylene oxide (100%) was added (2 changes for 30 min at room temperature) followed by the addition of propylene oxide: resin (48 g TLV Resin, 16 g VH1 Hardener, 36 g TLV Hardener VH2) solution at 3:1 ratio (1 h, room temperature) and 1:1 ratio (overnight with lids off, room temperature) respectively. A pure resin preparation (TAAB Low Viscosity Resin – medium recipe) was then added (2 changes, 2 h per change) prior to embedding in the oven at 60 °C for 48 h. Samples were sectioned with a Histo diamond knife at 1000nm (1micron) thick using Leica EM UC6 (Leica Biosystems, Wetzlar, Germany), placed on APES coated slide and stained using 1% Toluidine Blue in 1% Sodium Borate. Sections were imaged using ZEISS Axiophot microscope and captured using Micromanager Software. Cell area was measured by tracing individual cells taken from images taken at 200X magnification using ImageJ [47].

Elemental analysis

Sample digestates were diluted 1-in-10 using Milli-Q water prior to elemental analysis. The concentrations of 28 elements were obtained using inductively coupled plasma-mass spectrometry (ICP-MS; Thermo Fisher Scientific iCAPQ, Thermo Fisher Scientific, Bremen, Germany); Ag, Al, As, B, Ba, Ca, Cd, Cr, Co, Cs, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, P, Pb, Rb, S, Se, Sr, Ti, U, V, Zn. Operational modes included: (i) a helium collision-cell (He-cell) with kinetic energy discrimination to remove polyatomic interferences, (ii) standard mode (STD) in which the collision cell was evacuated, and (iii) a hydrogen collision-cell (H2-cell). Samples were introduced from an autosampler incorporating an ASXpress™ rapid uptake module (Cetac ASX-520, Teledyne Technologies Inc., Omaha, NE, USA) through a PEEK nebulizer

(Burgener Mira Mist, Mississauga, Burgener Research Inc., Canada). Internal standards were introduced to the sample stream on a separate line via the ASXpress unit and included Sc (20 µg L⁻¹), Rh (10 µg L⁻¹), Ge (10 µg L⁻¹) and Ir (5 µg L⁻¹) in 2% trace analysis grade HNO₃ (Fisher Scientific UK Ltd). External multi-element calibration standards (Claritas-PPT grade CLMS-2; SPEX Certiprep Inc., Metuchen, NJ, USA) included Ag, Al, As, B, Ba, Cd, Ca, Co, Cr, Cs, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, P, Pb, Rb, S, Se, Sr, Ti (semi-quant), U, V and Zn, in the range 0–100 µg L⁻¹ (0, 20, 40, 100 µg L⁻¹). A bespoke external multi-element calibration solution (Plasma-CAL, SCP Science, Courtaboeuf, France) was used to create Ca, K, Mg and Na standards in the range 0–30 mg L⁻¹. Boron, P and S calibration utilized in-house standard solutions (KH₂PO₄, K₂SO₄ and H₃BO₃). In-sample switching was used to measure B and P in STD mode, Se in H₂-cell mode and all other elements in He-cell mode. Sample processing was undertaken using Qtegra™ software (Thermo Fisher Scientific) with external cross-calibration between pulse-counting and analogue detector modes when required.

N measurements

Samples were measured using the Dumas method for N determination. The samples were dried for 17 h at 100 °C and then milled on a 1mm hammer mill. Prior to testing the sample were dried at 104 °C for 3 h and 1 g of sample was loaded on the instrument (Leco TruMacN Dumas gas analyser), following the manufacturer's instructions. Samples were converted to gases by heating in a combustion tube at 1150 °C. Interfering components are removed from the resulting gas mixture. The nitrogen compounds in the gas mixture or a representative part of the mixture, are converted to molecular nitrogen which is quantitatively determined by a thermal conductivity detector. The nitrogen content is then calculated by a microprocessor. To calculate the percent protein the standard ISO conversion factor of 5.7 was applied [48].

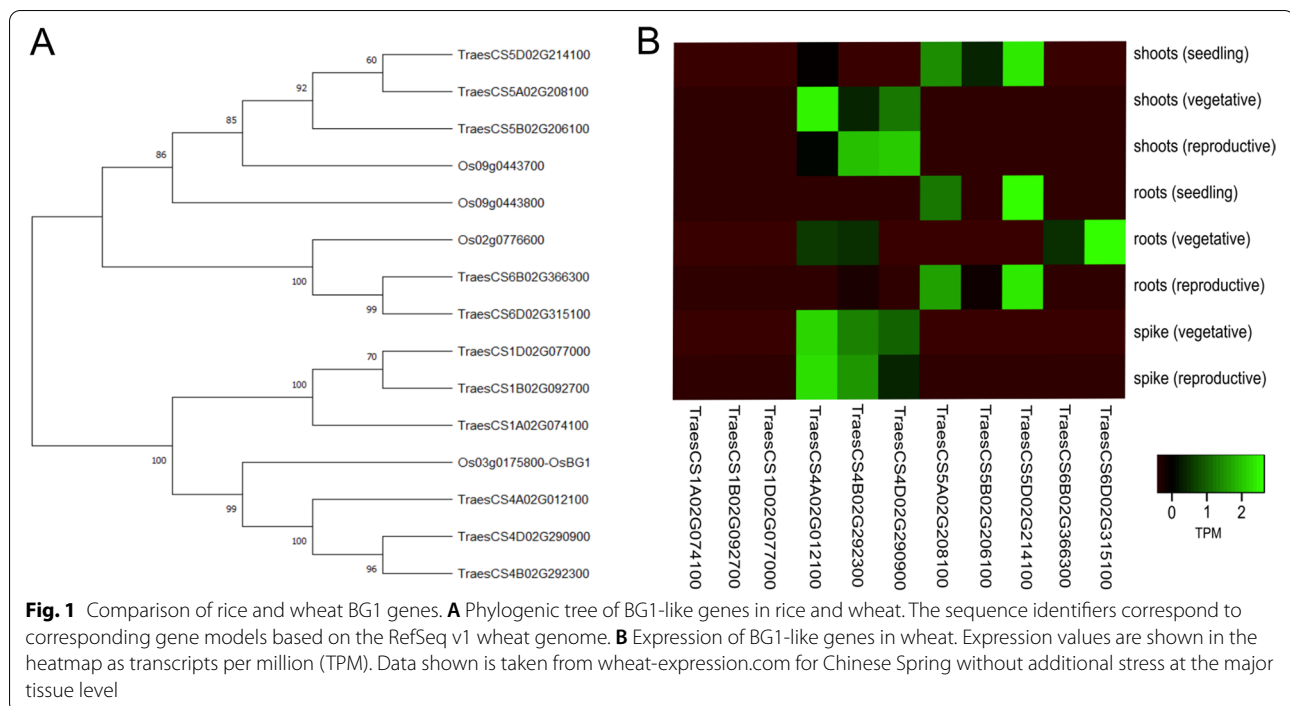
Statistical analyses

Analysis of variance (ANOVAs) or Wilcoxon Tests were run using the aov and TukeyHSD functions in the R environment with the null hypothesis of no difference between lines [49]. Tukey's post hoc test was added to identify each significant interaction between the lines tested. Data was plotted using R ggplot2 [50].

Results

Identification of BG1 orthologues in wheat

To identify *OsBG1* orthologues in wheat the amino acid sequence for Interpro ID IPR039621 was used in a BlastP



search against the Ensembl v 42 database. In rice there are four *BG1* like genes but in wheat eleven genes containing the big grain domain were identified (Fig. 1 A). Alignment of the rice and wheat *BG1* like proteins reveals two rice genes *Os09g0443700* and *Os090443800* cluster with a set of genes located on chromosome 5, a third rice gene (*Os02g077660*) clusters with a pair of genes on chromosome 6 and the final rice gene *OsBG1* (*Os03g0175800*) clusters with a set of gene on chromosome 1. The homoeologous gene triads which are closely related to *OsBG1* located on wheat chromosomes 1 and 4. However alignments of the predicted amino acid sequences suggest that the homoeologous genes on chromosome 4 share 73 to 76% identity whereas the chromosome 1 homoeologues shared 51–52% identity with *OsBG1* at the amino acid level (Suppl. Table 1). To further narrow down an ortholog of *OsBG1*, the expression profiles of the eleven *BG1*-like genes from wheat in public expression databases were examined. This indicated that the genes located on chromosome 4 are the most highly expressed of the family members. The homoeologous genes on chromosome 4 were expressed later in development and not during the seedling stage of growth and were the only genes to show expression in the spike tissues. This contrasts with the genes showing weaker expression located on chromosome group 5 mainly expressed during the seedling stage in both the roots and shoots but also in the root during reproduction/flowering. The 5B homoeologue does not appear to be expressed in any of

the main tissues. The homoeologues on chromosome group 6 were only expressed in the roots during the vegetative growth portion of the life cycle. While current gene models do not show a direct homoeologue on chromosome 6A, the lack of a homoeologous gene on 6A might be an artifact of the current gene model prediction as a gene model appeared to exist in previous versions of the genome under the gene model *TraesCS6A02G526800LC*, but is now not found in version Ref Seq v1.1 of the wheat genome. No observable expression of the genes located on chromosome 1 in the tissues was identified (Fig. 1B). It should also be noted that of the three sets of genes on chromosomes 4, 5 and 6 the homoeologous set of genes are expressed at different levels and sometimes in different tissues suggesting some type of differential regulation of the *BG* family in wheat. The unique expression of the chromosome 4 homoeologues in the spike tissue led to the further characterization of the putative orthologs on chromosome 4 for further study. As the homeologue on 4A is expressed the highest, based on publicly available expression, the homeologue of 4A was specifically tested for overexpression studies.

Overexpression of *TaBG1-A* in wheat

To understand if the genes on chromosome 4 could be the functional orthologues of *OsBG1* *TaBG1-A* was overexpressed in wheat from the constitutive promoter *OsActin*. A total of thirty-seven independent transgenic lines were regenerated and the number of T-DNAs

inserted determined in each T_0 plant measured by *nptII* copy number analysis. To limit the possible adverse effects of multiple insertions, nine single- and five two-copy T-DNA insertion lines were prioritized for *TaBG1* expression determination. Three independent single insertion lines with increased expression in the leaves of the T_0 plants from 4.9 to 6.3-fold over endogenous expression of *TaBG1* in the leaf were taken forward for further analysis of grain and agronomic phenotypes (Suppl. Figure 1). These T_0 lines (OE-1, OE-2 and OE-3) were then selfed and the T_1 plants assayed for T-DNA copy number to identify lines homozygous for the T-DNA insertion, which were then grown to T_2 seed for further analysis. Several agronomic traits were measured for differences between the overexpression lines and a null segregant hereto referred as WT. The three overexpression lines were grown on a low fertility soil with three different levels of N including a deficient (70 kg/ha equiv.), a moderate (140 kg/ha equiv.) and an adequate level of N (210 kg/ha equiv.) added to understand the role

N may play in seed size. Average seed weight did not vary with changes in N level in any of the four lines tested, however all three overexpression lines showed significantly increased TGW than a segregated null line used as a control (Fig. 2 A). Yield did significantly increase with increasing N supplied in a dose dependent manner (p val<0.01) but comparisons of the three overexpression lines to WT at each N level showed no significant differences (Fig. 2 C). When the overexpression lines were compared to control lines the increased TGW did not increase overall yields. The lack of an increase in yield was due to the significant decrease in the number of seeds per plant in the overexpression lines, (p val<0.05), suggesting that the plant simply traded off seed number to increase seed size (Fig. 2B). Other potential agronomic traits such as tiller number, which we defined as the tillers which produced an inflorescence, or biomass also did not differ significantly between the overexpression lines and WT plants (Fig. 2D and Suppl. Figure 2). The fact that there were no significant differences in biomass,

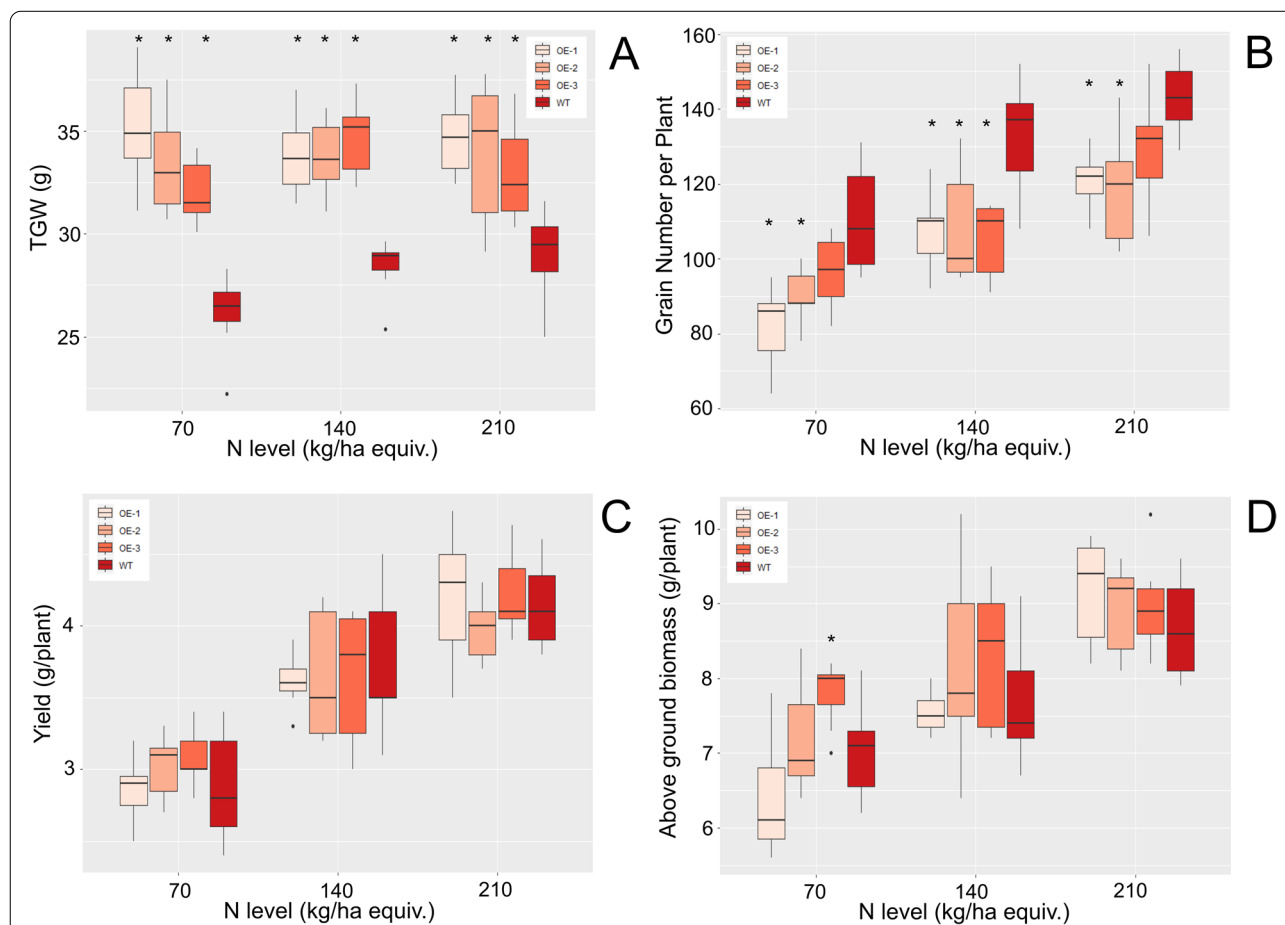


Fig. 2 Agronomic measurements of the overexpression of *TaBG1-A*. Shown are the measurements from twenty-one plants per treatment under three levels of nitrogen, low (70), sufficient (140), and high (210) levels of N. **A** Seed weight (TGW) **B** Grain number per plant **C** Yield per plant **D** Above ground biomass per plant. Significant differences are shown by * which represent a p val. < 0.05 relative to WT Fielder

apart from one line OE-3 grown on 70 kg/ha N, and no significant differences in yield, then led to no change to the harvest index for any treatment (p val OE-1=0.98, OE-2=0.24, OE-3=0.13) (Suppl. Figure 3).

Components of increased grain size

To understand which seed parameters were modified to increase the overall size of the wheat seeds the three OE lines were analyzed using MARVIN to measure the length and width components of grain size. From this it was found that on average the *TaBG1-A* overexpressing lines were marginally longer in length with only OE-2 showing significantly increased length (10% greater than WT). While the other lines also appeared to show an increased average seed length, 3% for both OE-1 and OE-3, the differences were not significant. Strikingly, seed width was significantly increased in the three OE lines with seeds of 10–26% greater width than WT seeds, p val<0.001 (Fig. 3). Cross sections of ten grains from line OE-1 and WT were also taken in two dimensions to understand the cause of these differences at a tissue level within the grain (Fig. 4). Differences were seen in the size of the aleurone layer cells with OE-1 having smaller and more numerous cells than WT grains (Fig. 5). OE-1 lines showed an average decrease of 21% cell area relative to WT aleurone cells.

Nutritional Content of *TaBG1* overexpression lines

To understand the effect larger grain size could have on nutritional traits such as mineral content, grains were analyzed by ICP-MS to measure the mineral content (Fig. 6 A-C). All three of the lines tested showed significantly lower Zn levels in the grain relative to WT (p val, <0.05, <0.001, <0.05). This contrasted with Fe and Mn levels as no significant differences were seen in these

micronutrients. Other essential element concentrations were also altered in *TaBG1-A* overexpression lines, but only in two of the three lines tested; these include Ca, K and P (Suppl. Figure 4). One line OE-1 was also significantly lower in S (p val<0.05). No differences were seen in any of the lines tested for Mg or B.

Nitrogen content of the grain was also determined from the plants grown under three different N fertilization levels using the Dumas method to measure the nitrogen content. As seen in Fig. 6D there were significantly lower levels of N in the grains of the OE lines versus WT when grown under a moderate level of N (140 kg/ha equiv.). There were no significant differences seen in the N content of the grain when grown on low or high levels of N. The lack of difference in yields also did not lead to any changes in nitrogen use efficiency (NUE) defined as yield divided by N concentration in the soil or nitrogen efficiency ratio (NER) defined as yield divided by tissue N concentration at a given N concentration in the soil [51]. Using the nitrogen efficiency ratio there was a general trend towards a higher NER as lower N content in the grain but the p values ranged from 0.1 to 0.8 (Suppl. Figure 5).

Discussion

A number of agronomic traits are believed to be involved increasing yield in plants. Grain size is one such trait and the target of numerous studies [14–27]. To understand the role in which increasing grain size can help to increase yields in wheat the orthologues of the rice Big Grain 1 (*OsBG1*) gene was identified in wheat and chosen for further characterization. *BG1* was chosen as previous reports had shown that *OsBG1* could improve yields 15–20% under field conditions with increased expression in rice [52]. The putative orthologues of *OsBG1* in



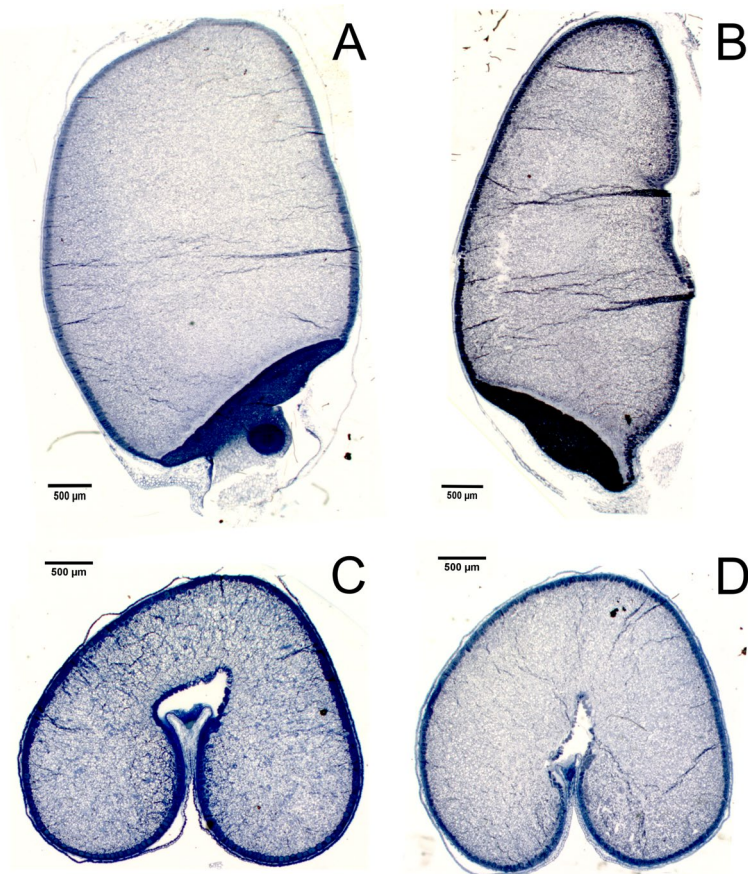


Fig. 4 Longitudinal and transverse sections of *TaBG1-A* (OE-1) overexpression line and WT Fielder. **A** OE-1 longitudinal section **B** WT longitudinal section **C** OE-1 transverse section **D** WT transverse section. Scale bar = 500μm

wheat identified on chromosome 4 were the only members of the 11 Big Grain like genes found in wheat to be expressed in the spike (Fig. 1). Increasing expression of the A homoeologue of *TaBG1* on chromosome 4 did in fact increase seed size with seeds on average 25% larger by weight than a null segregant WT control (Fig. 2). This increase in grain size was mostly attributed to an increase in width and not to seed length, as seed length was only marginally increased (Figs. 3 and 4). At a seed tissue level cells in the mature grain of OE-1 showed a smaller average size than that of WT cells (Fig. 5). This average aleurone cell size difference was 21% smaller in OE-1 grains relative to WT suggesting that the increase in size of the grain is due to increased endosperm cell number and/or cell size. However, the increase in seed size did not drive an increase in yield as the larger seed size was offset by the lower seed number per plant, to maintain yield levels (Fig. 2). This is expected as similar results were seen in rice using the rice actin promoter and only when *OsBG1* was driven by its own promoter were biomass and yields increased [52]. Other recent attempts to increase yield in

wheat by increasing grain size have also been hampered by the negative association between seed size and seed number [53–55]. Even efforts by recurrent selection for higher seed size in wheat breeding programs showed that increases of up to 32% in seed size were possible but this increase was completely offset by reductions in seed number keeping yields static [56]. However one more recent transgenic approach has been able to overcome this problem by increased expression of an α -expansin from a grain specific promoter [57]. It is interesting that without the increased biomass an increase in overall yield is not seen which may suggest that seed size *per se* is not the most limiting yield factor and that during flowering the plant can switch between alternative pathways which affect seed number and yields. Further investigation into the spatial and temporal expression patterns of both *TaBG1* and *OsBG1* may help to elucidate how small changes in auxin signaling promote growth and ultimately yield.

As N is a major driver of increased yields and does lead to increased biomass, any increase in yields

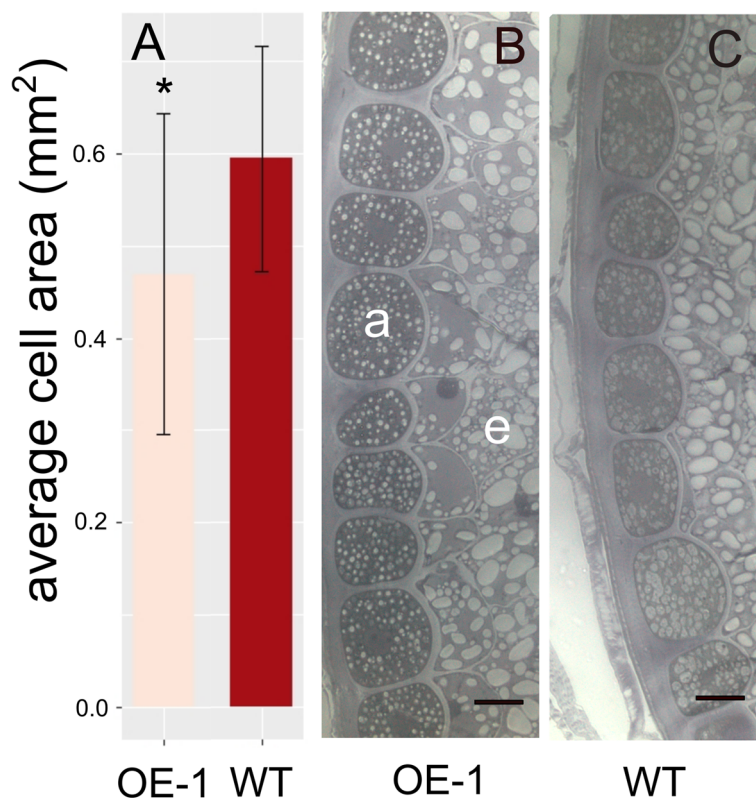


Fig. 5 Comparison of cell size in the fully developed grain. **A** average aleurone cell size in OE-1 versus WT grains. Asterisk represents a significant difference (p val. 0.03). **B** and **C** representative images of OE-1 (**B**) and WT (**C**) aleurone and endosperm cells at 200X magnification. Scale bar = 0.1 mm, a signifies an aleurone cell and e signifies endosperm cells

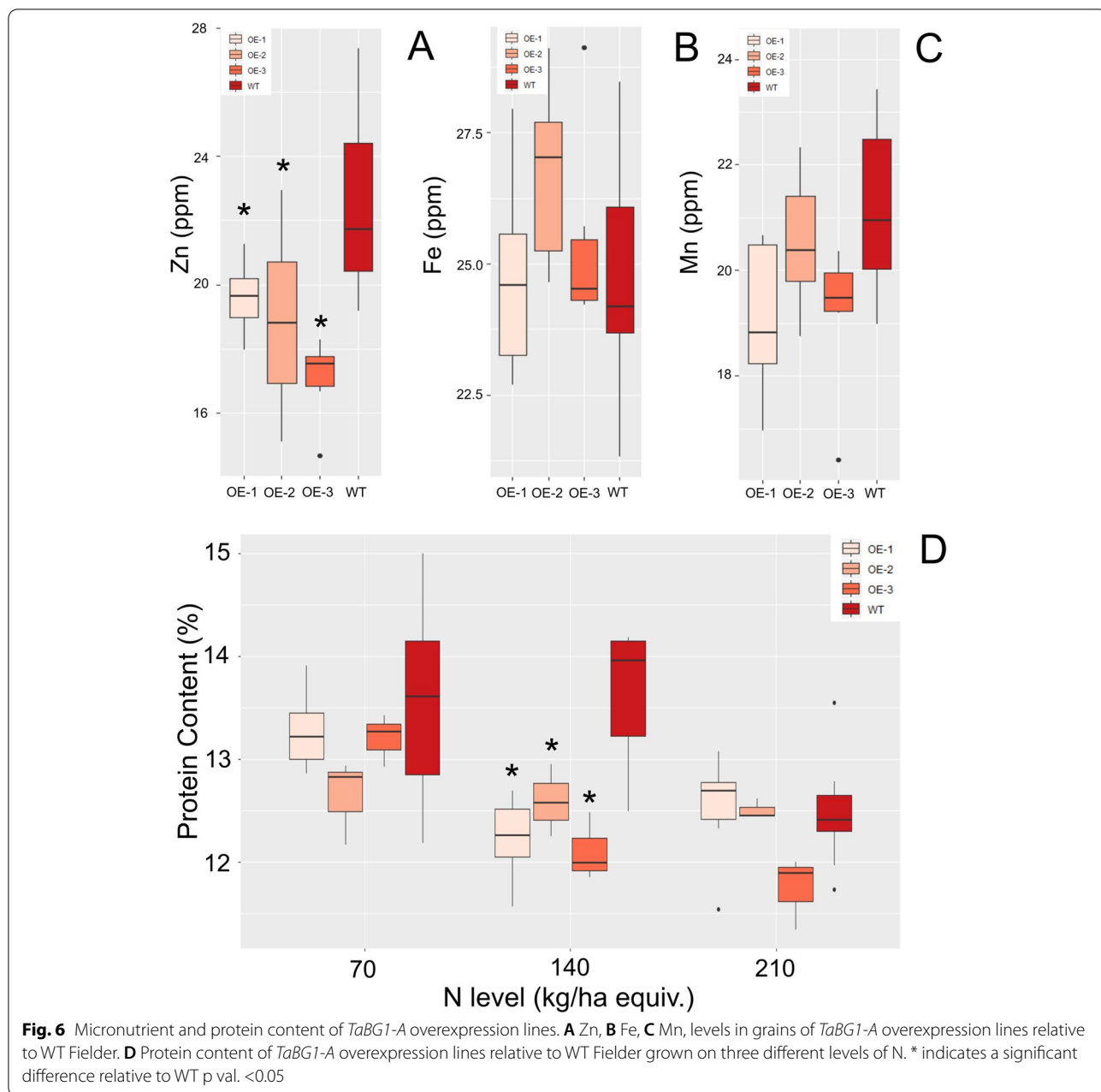
without an alteration in nitrogen supplied would also increase NUE further driving agronomic improvement [58]. Interestingly no change in the size of the grains were seen in the seeds with increased N treatments. WT size did seem to trend toward larger seeds, but the differences were not significant among the treatments tested and might have been partially offset by the increased seed number with increasing N levels supplied (Fig. 2). To understand the interaction between increase size grain NUE we set out to study the effects that larger grains could have on both the protein content and elemental concentrations of larger grains. It was found that under moderate N levels increasing grain size significantly decreased the protein content of the grain. This was the only level of N for which a difference could be seen and may suggest that increased N is again more important in helping to drive biomass gains leading to increased yields and more specific N treatments are necessary to understand the overall effects of protein content. This could also lead to lower NUE as yields are not increased and overall N in the grain is lower. This did not unfortunately lead to a higher NER which would sacrifice N in the grain to increase yields

however this was not seen in any of the overexpression lines.

The increased seed size also led to significantly lower Zn levels in the grain but not other micronutrients such as Fe or Mn. It has been previously shown that altered auxin levels can lead to lower Zn levels [59, 60]. One hypothesis is that this is due to the lower ratio of aleurone cells to endosperm cells resulting in a different in micronutrient density [36]. To a large extent similar patterns were identified here as P, Zn and K were lower in most of the over expression lines. This is most likely explained by the distribution in the wheat grain as Zn and P are in the aleurone layer but elements such as K are stored in the embryo. Other elements more widely distributed like S and Mg showed no clear trends suggesting that lower tissue types can explain the variation in most elements except Fe [61].

Conclusions

We have identified the orthologues of BG1 genes in wheat and demonstrated that alteration in expression of one of these genes can lead to significantly larger



grains. However constitutive expression of *TaBG1* alone will not increase overall yield as the plant compensates for the larger seed size by producing fewer grains. Other tradeoffs from this increased seed size are the change in the nutritional profile of the seeds, with lowered Zn and P levels but not Fe, S or Mg and potentially lower protein content under moderate application of N. Manipulation of *TaBG1* alone increased seed size significantly. However, this increase in seed size is not sufficient to increase yields in wheat but could be

a target for further breeding efforts in combination with other genes for increased biomass. Further work to evaluate whether the expression of *TaBG1* or *OsBG1* expressed from their native promoters may compensate for the reduced grain number to increase yield in wheat should also be considered.

Abbreviations

BG1: Big Grain 1; N: Nitrogen; Zn: Zinc; P: Phosphate; Fe: Iron; S: Sulphur; K: Potassium; Mn: Manganese; WT: Wild Type; NUE: Nitrogen Use Efficiency; NER: Nitrogen Efficiency Ratio; Ca: Calcium; Mg: Magnesium; B: Boron; TGW: Thousand Grain Weight; cDNA: complementary DNA.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-03294-x>.

Additional file 1.

Additional file 2.

Acknowledgements

The authors would like to thank The University of Nottingham school of Chemistry for running the ICP-MS samples for elemental content. Helen Appleyard and Marianna Rizzo at NIAB analytical services for analysis for N content analysis of the grains. Grain cross section work was performed at the School of Life Sciences Imaging, University of Nottingham, by Tim Self and Denise Mclean.

Authors' contributions

MM & EW designed the experiments. MM, MC and SB performed the experiments. The manuscript was written by MM and revised by EW. All authors read and approved the final manuscript.

Funding

This work was supported by the UK Biotechnology and Biological Sciences Research Council grant BB/N013441/1. The funding body played no role in the design of the study and collection, analysis, and interpretation of data, or in the writing of the manuscript.

Availability of data and materials

Data and seed are available upon request from the corresponding author. Seed materials will be transferred under MTA.

Declarations

Ethics approval and consent to participate

Experimental research and field studies on plants complies with relevant institutional, national, and international guidelines and legislation.

Consent for publication

All authors consent to publication.

Competing interests

The authors declare no competing interests.

Received: 13 May 2021 Accepted: 26 October 2021

Published online: 10 November 2021

References

- www.fao.com. FAOSTAT [Internet]. 2017. Available from: <http://www.fao.org/faostat/en/#home>
- Knight S, Kightley S, Bingham I, Hoad S, Lang B, Philpott H, et al. Project Report No. 502. Desk study to evaluate contributory causes of the current 'yield plateau' in wheat and oilseed rape. 2012;226.
- Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, Wright J, et al. An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res*. 2017;27:885–96.
- International Wheat Genome Sequencing Consortium (IWGSC) TIWGSC, IWGSC RefSeq principal investigators: R, Appels R, Eversole K, Feuillet C, Keller B, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 2018;361:eaar7191.
- Zimin A V, Puiu D, Hall R, Kingan S, Clavijo BJ, Salzberg SL. The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *Gigascience*. 2017;6:1–7.
- Mayer KFXX, Rogers J, Dole el J, Pozniak C, Eversole K, Feuillet C, et al. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* (80-). 2014;345:1251788.
- Würschum T, Leiser WL, Langer SM, Tucker MR, Longin CFH. Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theor Appl Genet*. 2018;131:2071–84.
- Jamieson PD, Martin RJ, Francis GS. Drought influences on grain yield of barley, wheat, and maize. *New Zeal J Crop Hortic Sci*. 1995;23:55–66.
- Fischer RA, Aguilar I, Laing DR. Post-anthesis sink size in a high-yielding dwarf wheat: Yield response to grain number. *Aust J Agric Res*. 1977;28:165–75.
- Yoshida H, Horie T, Shiraiwa T. A model explaining genotypic and environmental variation of rice spikelet number per unit area measured by cross-locality experiments in Asia. *F Crop Res*. 2006;97:337–43.
- Yoshida S. Fundamentals of rice crop science. 1981.
- Zhang X, Wang J, Huang J, Lan H, Wang C, Yin C, et al. Rare allele of OsP-PKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc Natl Acad Sci U S A*. 2012;109:21534–9.
- Zuo J, Li J. Molecular Genetic Dissection of Quantitative Trait Loci Regulating Rice Grain Size. *Annu Rev Genet*. 2014;48:99–118.
- Ma D, Yan J, He Z, Wu L, Xia X. Characterization of a cell wall invertase gene TaCwi-A1 on common wheat chromosome 2A and development of functional markers. *Mol Breed*. 2012;29:43–52.
- Jiang Y, Jiang Q, Hao C, Hou J, Wang L, Zhang H, et al. A yield-associated gene TaCWI, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theor Appl Genet*. 2015;128:131–43.
- Hanif M, Gao F, Liu J, Wen W, Zhang Y, Rasheed A, et al. TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. *Mol Breed*. 2016;36:1–8.
- Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N, Onodera H, et al. Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat Genet*. 2013;45:707–11.
- Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, del Blanco A, et al. 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*. Springer Verlag; 2016;129:1099–112.
- Zhang Y, Li D, Zhang D, Zhao X, Cao X, Dong L, et al. Analysis of the functions of TaGW2 homoeologs in wheat grain weight and protein content traits. *Plant J*. 2018;94:857–66.
- Zhai H, Feng Z, Du X, Song Y, Liu X, Qi Z, et al. 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*. 2018;131:539–53.
- Guo Y, Sun J, Zhang G, Wang Y, Kong F, Zhao Y, et al. Haplotype, molecular marker and phenotype effects associated with mineral nutrient and grain size traits of TaGS1a in wheat. *F Crop Res*. 2013;154:119–25.
- Ma L, Li T, Hao C, Wang Y, Chen X, Zhang X. TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnol J*. 2016;14:1269–80.
- Dong L, Wang F, Liu T, Dong Z, Li A, Jing R, et al. Natural variation of TaGASR7-A1 affects grain length in common wheat under multiple cultivation conditions. *Mol Breed*. Kluwer Academic Publishers; 2014;34:937–47.
- Milner MJ, Howells RM, Craze M, Bowden S, Graham N, Wallington EJ. A PSTOL-like gene, TaPSTOL, controls a number of agronomically important traits in wheat. *BMC Plant Biol*. 2018;18:115.
- Sajjad M, Ma X, Habibullah Khan S, Shoaib M, Song Y, Yang W, et al. TaFlo2-A1, an ortholog of rice Flo2, is associated with thousand grain weight in bread wheat (*Triticum aestivum* L.). *BMC Plant Biol*. 2017;17:164.
- Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, et al. Expression of TaCYP78A3, a gene encoding cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. *Plant J*. Blackwell Publishing Ltd; 2015;83:312–25.
- Ma M, Zhao H, Li Z, Hu S, Song W, Liu X. TaCYP78A5 regulates seed size in wheat (*Triticum aestivum*). *J Exp Bot*. 2016;67:1397–410.
- Mora-Ramirez I, Weichert H, Wirén N von, Froberg C, Bodt S de, Schmidt R-C, et al. The da1 mutation in wheat increases grain size under ambient and elevated CO2 but not grain yield due to trade-off between grain size

- and grain number. *Plant-Environment Interact.* John Wiley & Sons, Ltd; 2021;2:61–73.
29. Sinclair TR, Jamieson PD. Grain number, wheat yield, and bottling beer: An analysis. *F Crop Res.* Elsevier; 2006;98:60–7.
 30. Makino A. Photosynthesis, grain yield, and nitrogen utilization in rice and wheat. *Plant Physiol.* American Society of Plant Biologists; 2011;155:125–9.
 31. Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, et al. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor Appl Genet.* 2006;112:1164–71.
 32. Mao H, Sun S, Yao J, Wang C, Yu S, Xu C, et al. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc Natl Acad Sci U S A.* National Academy of Sciences; 2010;107:19579–84.
 33. Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, et al. Natural variation at the DEP1 locus enhances grain yield in rice. *Nat Genet.* 2009;41:494–7.
 34. Hu J, Wang Y, Fang Y, Zeng L, Xu J, Yu H, et al. A rare allele of GS2 enhances grain size and grain yield in rice. *Mol Plant.* Cell Press; 2015;8:1455–65.
 35. Public Health England. National Diet and Nutrition Survey: results from Years 1 to 4 (combined) of the rolling programme for 2008 and 2009 to 2011 and 2012 - Publications - GOV.UK.
 36. Fan M-S, Zhao F-J, Fairweather-Tait SJ, Poulton PR, Dunham SJ, McGrath SP. Evidence of decreasing mineral density in wheat grain over the last 160 years. *J Trace Elem Med Biol.* 2008;22:315–24.
 37. Bates B, Lennox A, Prentice A, Bates CJ, Page P, Nicholson S, et al. National Diet and Nutrition Survey: Results from Years 1, 2, 3 and 4 (combined) of the Rolling Programme (2008/2009–2011/2012): A survey carried out on behalf of Public Health England and the Food Standards Agency. Public Health England; 2014.
 38. Shewry PR, Hey SJ. The contribution of wheat to human diet and health. *Food Energy Secur.* Wiley-Blackwell Publishing Ltd; 2015. p. 178–202.
 39. Vogel KP, Johnson VA, Mattern PJ. Protein and Lysine Content of Grain, Endosperm, and Bran of Wheats from the USDA World Wheat Collection 1. *Crop Sci.* Wiley; 1976;16:655–60.
 40. Borrill P, Ramirez-Gonzalez R, Uauy C. expVIP: A customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* American Society of Plant Biologists; 2016;170:2172–86.
 41. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms.
 42. Babicki Sasha, Arndt David, Marcu Ana, Liang Yongjie, Grant Jason R, Maciejewski Adam, et al. Heatmapper: web-enabled heat mapping for all | *Nucleic Acids Research | Oxford Academic.* *Nucleic Acids Res.* 2016;44:W147–53.
 43. Ishida Y, Tsunashima M, Hiei Y, Komari T. Wheat (*Triticum aestivum* L.) Transformation Using Immature Embryos. *Methods Mol Biol.* 2015. p. 189–98.
 44. Risacher T, Craze M, Bowden S, Paul W, Barsby T. Highly Efficient Agrobacterium-Mediated Transformation of Wheat Via In Planta Inoculation. *Methods Mol Biol.* 2009. p. 115–24.
 45. Bates R, Craze M, Wallington EJ. Agrobacterium-Mediated Transformation of Oilseed Rape (*Brassica napus*). *Curr Protoc Plant Biol.* Hoboken, NJ, USA; 2017. p. 287–98.
 46. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods.* 2001;25:402–8.
 47. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods.* Nature Publishing Group; 2012. p. 676–82.
 48. ISO 16634-2:2016(en), Food products — Determination of the total nitrogen content by combustion according to the Dumas principle and calculation of the crude protein content — Part 2: Cereals, pulses and milled cereal products [Internet]. [cited 2021 Jul 12]. Available from: <https://www.iso.org/obp/ui/#iso:std:iso:16634:-2:ed-1:v1:en>
 49. Computing RF for S. R Core Team (2018). R: A language and environment for statistical computing [Internet]. 2018. Available from: <https://www.r-project.org/>
 50. Wickham H. Introduction. *ggplot2.* New York, NY: Springer New York; 2009.
 51. Hammond JP, Broadley MR, White PJ, King GJ, Bowen HC, Hayden R, et al. Shoot yield drives phosphorus use efficiency in *Brassica oleracea* and correlates with root architecture traits. *J Exp Bot.* 2009;60:1953–68.
 52. Liu L, Tong H, Xiao Y, Che R, Xu F, Hu B, et al. Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. *Proc Natl Acad Sci U S A.* 2015;112:11102–7.
 53. Bustos DV, Hasan AK, Reynolds MP, Calderini DF. Combining high grain number and weight through a DH-population to improve grain yield potential of wheat in high-yielding environments. *F Crop Res.* 2013;145:106–15.
 54. Quintero A, Molero G, Reynolds MP, Calderini DF. Trade-off between grain weight and grain number in wheat depends on GxE interaction: A case study of an elite CIMMYT panel (CIMCOG). *Eur J Agron.* 2018;92:17–29.
 55. Molero G, Joynson R, Pinera-Chavez FJ, Gardiner L-J, Rivera-Amado C, Hall A, et al. Elucidating the genetic basis of biomass accumulation and radiation use efficiency in spring wheat and its role in yield potential. *Plant Biotechnol J.* 2019;17:1276–88.
 56. Wiersma JJ, Busch RH, Fulcher GG, Harelend GA. Recurrent Selection for Kernel Weight in Spring Wheat. *Crop Sci.* 2001;41:999–1005.
 57. Calderini DF, Castillo FM, Arenas-M A, Molero G, Reynolds MP, Craze M, et al. Overcoming the trade-off between grain weight and number in wheat by the ectopic expression of expansin in developing seeds leads to increased yield potential. *New Phytol.* 2021;230:629–40.
 58. Zörb C, Ludewig U, Hawkesford MJ. Perspective on Wheat Yield and Quality with Reduced Nitrogen Supply. *Trends Plant Sci.* 2018;23:1029–37.
 59. Skoog F. Relationships Between Zinc and Auxin in the Growth of Higher Plants. *Am J Bot.* 1940;27:939–51.
 60. TSUI C. The role of zinc auxin synthesis in the tomato plant. *Am J Bot.* 1948;35:172–9.
 61. De Brier N, Gomand S V, Donner E, Paterson D, Smolders E, Delcour JA, et al. Element distribution and iron speciation in mature wheat grains (*Triticum aestivum* L.) using synchrotron X-ray fluorescence microscopy mapping and X-ray absorption near-edge structure (XANES) imaging. *Plant Cell Environ.* 2016;39:1835–47.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

