New Phytologist

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

Daniel F. Calderini[1]*, Francisca M. Castillo[1,2], Anita Arenas-M[1,2], Gemma Molero[3], Matthew P. Reynolds[3], Melanie Craze[4], Sarah Bowden[4], Matthew J. Milner[4], Emma J. Wallington[4], Adam Dowle[5], Leonardo D. Gomez[5] and Simon J. McQueen-Mason[5]*

[1] Institute of Plant Production and Protection, Universidad Austral de Chile, Campus Isla Teja, 5090000, Valdivia, Chile.

[2] Institute of Biochemistry and Microbiology, Faculty of Sciences, Universidad Austral de Chile, 5090000, Valdivia, Chile.

[3] International Maize and Wheat Improvement Center (CIMMYT). El Batán, Texcoco. CP 56237, Mexico.

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

[5] CNAP, Biology Department, University of York, Wentworth Way, Heslington, York YO10 5YW, UK.

Acceptance date: 21 October 2020

ORCID information:

Daniel F Calderini	0000-0001-7873-5101
Francisca M. Castillo	0000-0001-9039-5736
Anita Arenas-M	0000-0002-8459-164X
Emma Molero	0000-0002-6431-7563
Matthew P Reynolds	0000-0002-4291-4316
Melanie Craze	0000-0003-0757-0914
Sarah Bowden	0000-0001-5105-076X
Matthew J. Milner	0000-0001-9184-7423
Emma J. Wallington	0000-0003-3715-7901
Adam Dowle	0000-0002-6501-5444
Leonardo D. Gomez	0000-0001-6382-9447
Simon J. McQueen-Mason	0000-0002-6781-4768

*Author for correspondence:
Daniel F. Calderini *Tel: +56* 632221723 *Email:* danielcalderini@uach.cl
Simon J. McQueen-Mason *Tel: +44 1904328775* *Email:* simon.mcqueenmason@york.ac.uk

Total word count (excluding summary, references and legends):	4828	Nº of figures:	7
Summary:	155	Nº of Tables:	2
Introduction:	839	Nº of Supporting Information (SI) files:	1 file with 3 figures and 4 tables
Materials and Methods:	1512		
Results:	1100		
Discussion:	1095		
Acknowledgements:	151		

Abstract

•Wheat is the most widely grown crop globally, providing 20 % of all human calories and protein. Achieving step changes in genetic yield potential is crucial to ensure food security, but efforts are thwarted by an apparent trade-off between grain size and number. Expansins are proteins that play important roles in plant growth by enhancing stress relaxation in the cell wall, which constrains cell expansion.

•Here, we describe how targeted over-expression of an α-expansin in early developing wheat seeds leads to a significant increase in grain size without a negative effect on grain number, resulting in a yield boost under field conditions.

•The best-performing transgenic line yielded 12.3 % higher average grain weight than the control, and this translated to an increase in grain yield of 11.3 % in field experiments using an agronomically appropriate plant density.

•This targeted transgenic approach provides an opportunity to overcome a common bottleneck to yield improvement across many crops.

***Keywords*:** expansin protein, grain weight, trade-off, pericarp, grain number, grasses, transgenic

Introduction

Increasing wheat yield is a global priority for food security (Foulkes *et al.*, 2011) since this crop provides ~20 % of calories and protein in human diets. However, rates of genetic gains in grain yield (GY) potential have decreased since the Green Revolution and further GY improvement requires new approaches (Foulkes *et al.*, 2011; Molero *et al.*, 2019). In the past, GY has been consistently increased by higher grain number per unit area (GN), however, the trade-off between average grain weight (GW) and GN has become a bottleneck for improving GY as demonstrated by recent studies in a wide range of wheat genotypes including elite materials from the International Maize and Wheat Improvement Center (CIMMYT) (Quintero *et al.*, 2018; Molero *et al.*, 2019; Rivera-Amado *et al.*, 2019).

Increasing individual GW has the potential to improve wheat yield, however, attempts to increase GY by increasing grain size have been hampered by the negative association

between GW and GN (Bustos *et al.*, 2013; Quintero *et al.*, 2018; Molero *et al.*, 2019). For example, recurrent selection for higher GW in wheat breeding programs showed that increases of up to 32 % in this trait were completely offset by reductions in GN (Wiersma *et al.*, 2001). More recently, genetic studies have been used to identify quantitative trait loci (QTL) associated with grain size in wheat and other crops (e.g. Gross *et al.*, 2003; Simmonds *et al.*, 2014; Griffiths *et al.*, 2015). For example, Brinton *et al.*, (2017) showed a 6.9 % increase in GW in near isogenic lines of wheat (NIL) carrying a QTL on chromosome 5A affecting grain size; and Wang *et al.*, 2018 reached the highest increase of GW in triple mutant lines of the *TaGW2* gene (≈20 %). However, in these and other cases increases in GW had little impact on GY due to the trade-off between GW and GN (Song *et al.*, 2007; Zhang *et al.*, 2018). Remarkably, this trade-off does not appear to be due to photo-assimilate restriction after anthesis (Slafer and Savin, 1994; Borrás *et al.*, 2004; Reynolds *et al.*, 2009; Quintero *et al*., 2018). Indeed, in an experiment using doubled haploid wheat lines from a cross designed to complement high grain number with large seed size, radiation use efficiency increased in response to sink strength during grain filling in comparison with the parental cultivars (Bustos *et al*., 2013). This indicates that breaking the negative association between GN and GY does not require increased photosynthesis during grain growth to increase GY. As source limitation cannot account for the apparent trade-off between GW and GN, other limiting factors must be at play.

The number and potential weight of grains are determined early in development, in a period starting just before anthesis and continuing until just beyond grain setting (Calderini *et al.*, 1999, 2001; Ugarte *et al.*, 2007; Brinton & Uauy, 2019). Grain expansion and grain filling are separated in time, with the ovary expansion preceding grain filling (Calderini *et al.*, 1999; Brinton & Uauy, 2019). Grain growth (i.e. grain enlargement and dry matter accumulation) involves the coordinated expansion of the maternally-derived pericarp and the seed endosperm, and is almost complete at the time when storage deposition begins (Fig. 1). The expansion of these tissues determines the holding capacity of the seed for storage reserves and therefore the potential final GW (Brinton and Uauy; 2019; Herrera and Calderini, 2020). Therefore, it seems possible that genetic manipulation of grain expansion

in a spatiotemporal manner, could increase the rate or duration of this growth phase in grain development and lead to increased yield potential.

In plants, cells grow through turgor-driven expansion that is constrained by the cell wall (Cosgrove, 2018). Cell wall extensibility is under dynamic control in plant cells, and expansins play a key role by inducing the relaxation of the stress that is generated in the cell wall through the action of turgor pressure (McQueen-Mason *et al.*, 1992). Expansin manipulation can lead to changes in growth and development (Fleming *et al.*, 1997; Pien *et al.*, 2001; Rochange *et al.*, 2001; Choi *et al.*, 2003). These cell wall proteins appear to act by disrupting non-covalent associations between cellulose and matrix polysaccharides in the plant cell wall, allowing the polymers to slip relative to one another, relaxing stress in the wall and allowing it to extend (McQueen-Mason & Cosgrove, 1994). Expansins fall into two well-separated groups, designated as α- and β-expansins, based on sequence homology and activity; with α-expansins more clearly associated with cell expansion and growth (Cosgrove, 2018).

We previously revealed an association between grain expansion and the expression of α-expansins in wheat grain, and suggested that the expansin manipulation might provide a way to increase GW (Lizana *et al.*, 2010). In the work presented here, we increased the amount of α-expansin in developing grain, by the ectopic expression of *TaExpA6,* (an expansin gene normally expressed in wheat roots) under control of a grain specific gene promoter; and show that this can lead to increase GW without a negative impact on GN and, in turn to improve GY (Fig. 1).

Materials and Methods

Genetic constructs assembly

For this work we transformed wheat with a construct containing the *TaExpA6* coding sequence under the control of wheat *puroindoline-b (PinB)* gene promoter, which drives transcription restricted to the endosperm, aleurone and pericarp layers in developing seeds, but not in the embryo (Gautier et al., 1994; Digeon et al., 1999). We verified that the expression of the *TaPinB* gene is specific to the mentioned tissues using the wheat expression browser (Ramírez-González, et al., 2018). The intermediate cassette containing

the *TaPinb*-promoter (RefSeq v1.1: *TraesCS7B02G431200*) controlling the *TaExpA6* CDS was recombined into the binary vector pRLF12-R1R2 to create pEW279-Exp, using a Gateway LR Clonase II kit (Thermo Fisher Scientific Inc.). Figure S1 shows the T-DNA region of this construct. The plasmid was electro-transformed into *Agrobacterium tumefaciens* strain EHA105 (Hood *et al.*, 1993). The sequence of *TaExpA6* gene was obtained from GeneBank accession "*AY543532*" (www.ncbi.nlm.nih.gov/nuccore/AY543532) and verified in the recently released genome sequence of Chinese Spring variety, RefSeq v1.1(*TraesCS4A02G034200*) (Appels *et al.*, 2018).

Wheat transformation and subsequent wheat generations

Transformation of immature embryos isolated from spring wheat cultivar 'Fielder' was carried out by co-cultivation of pre-treated immature wheat embryos with *Agrobacterium* containing pEW279-Exp at 23 ºC in the dark for 2 days (Ishida *et al.,* 2014). Following removal of the embryonic axis, tissue culture was performed essentially as described previously (Risacher *et al.,* 2009). From thirty-seven regenerated wheat plants, DNA was isolated using DNA extraction protocol describe by Howells *et al*. (2018), and the plants were confirmed as transformant by PCR amplification of the transgene using primers Exp6-Forward (5`-CCG TTC TCG CGT TCT GCT TCGT-3`) and NosT-reverse (5`-CGA TCG GGT GAA ATT CGG ATCC-3`) using FastStart Taq polymerase (Thermo Fisher Scientific Inc.) with an annealing temperature of 53 ºC. A transformation efficiency of 30.6 % was achieved with this construct, calculated as the percentage of wheat embryos from which a transformed wheat plant was regenerated. It is important to point out that the control wheat plants used in the experiments correspond to cultivar Fielder that underwent the same tissue culture process as the transformed lines.

From each wheat transformant plant (T0), 10 individuals plants were grown in pots in a greenhouse (10 individuals T1 plants from 37 lines, total 370 plants) at The University of York and the presence of *TaExpA6* was checked by PCR using a forward primer Ta-*PinB*-F (5`-ACAACACACAATGGTAGGCAAA-3) and reverse primer *TaExpA6*-R (5`-GGTCCCCTTCACCGACAT-3`). T-DNA copy number was determined by qPCR assay of

the *nptII* gene relative to a single copy wheat gene amplicon, normalized to a known single copy wheat line in T0 and T1 plants (Table S1). In the next generations (T2, T3, T4) we carried out genomic DNA extraction from leaves using the CTAB method standardized protocol and copy number determination was performed by IDna Genetics services (UK) (https://www.idnagenetics.com/) (Table S2).

The multiplication of T2 seed was also carried out in greenhouses in 2015/16. During the vegetative stage, leaves were sampled for PCR analysis to identify homozygous lines. Lack of negative segregants among the sampled progeny (at least ten individuals sampled per line) was used as an indicator of homozygosis. Twenty four homozygous lines were selected to be evaluated in field experiments (T3 and T4 generations).

Experiments and field conditions (T3 and T4 generations)

Two field experiments were carried out at the Experimental Station of the Universidad Austral de Chile in Valdivia (39°47'S, 73°14'W). Experiment 1 was a low plant density (LPD) experiment with T3 lines during the 2016/17 growing season. The aim of the low plant density experiment was to increase the seed bank and evaluate the performance of lines under field conditions. Experiment 2 was sown at regular plant density (RPD) with four selected T4 lines in the 2018/19 season.

In the LPD experiment, 24 wheat transformed lines and the control spring cultivar Fielder were sown at a rate of 44 pl m^{-2}. Plots were 1.5 m long and 0.6 m wide, with 5 rows 0.15 m apart and 0.15 m between seeds in each row. In the RPD experiment four selected homozygous transformant lines and the control were sown at farmer´s conventional plant rate of 300 pl m^{-2}. In this experiments, the same plot dimensions and row spacing was used, but seed were sown at 0.022 m intervals. In both experiments additional rows were sown rounding each plot to avoid border effects. The experiments were arranged in a complete randomized block design with three replications. All plots were subjected to optimal agronomical management to assure high potential yield conditions. To this end, plots were fertilized, drip irrigation was applied to avoid water stress and weeds; and insects and diseases were prevented or controlled in both experiments.

In the RPD experiment, the four transformant lines were selected based on performance in the LPD experiment, low construct copy numbers, levels of transgene expression recorded in the previous generation (Table S1).

Crop phenology and measurements

Crop phenology was recorded following the decimal code scale (Zadoks *et al.*, 1974) at each plot. At harvest a 1 m length of the central row was sampled in each plot to determine grain yield and components in both experiments. Plant samples were oven-dried at 60°C for 48 h for dry weight. Grain number and average grain weight were measured. Grain yield per plant was calculated in the LPD and grain yield per square meter in RPD experiments. Grain dimensions (length, width, and area) were recorded using a Marvin seed analyzer (Wittenburg, Germany) (https://www.marvitech.de/en/home-2/) after grains were oven-dried at 60°C for 48 h.

RNA extraction and expression analyses by RT-qPCR

Grains of grain position one and two (G1, G2) from central spikelets of two main stem spikes (eight grains in total each replicate) were sampled at 5, 15 and 25 days after anthesis (DAA) in LPD experiment and at 5, 10, 15, 20 and 25 DAA in the RPD experiment for RNA extraction and gene expression analysis. Grains were sampled and frozen in liquid nitrogen. Total RNA was extracted using NucleoSpin™ columns (Macherey-Nagel) following the manufacture's protocol and standardizing the RNA extraction protocol based on Sangha *et al.*, (2010).

The *TaExpA6* expression in grains was assessed by quantitative reverse transcription PCR (RT-qPCR). cDNA was synthesized from 500 ng RNA (pretreated with DNaseI (Invitrogen) using ImProm-II™ Reverse Transcription System). The qPCR reactions with a final volume of 25 µL, was performed using the Brilliant II SYBR Green PCR Master Mix (Stratagene, Agilent technologies), 0.2 µM and primers Trangene*TaExpA6*_F1 (5`-ATCTCCACCACCACCAAAACA-3`) and Transgene*TaExpA6*_R1 (5`-GAAGCAGAACGCGAGAACGG-3`). No-template and no-transcriptase controls were included to detect genomic DNA contamination. The transcript abundance of the *TaExpA6*

gene in grains was determined using the method of *Pfaffl* (2001), where the ubiquitin-conjugating enzyme (*TraesCS4A02G414200)* gene was used as an internal control, using the primers *Forward* (5`- CGGGCCCGAAGAGAGTCT-3`) and *Reverse* (5`-ATTAACGAAACCAATCGACGGA-3`). Fluorescence raw data was analysed with LinRegPCR software for quantification analysis of gene expression (Ruijter *et al.*, 2009).

Protein Extraction and LC-MS proteomics

We extracted total proteins from wheat grains of the position two of the central spikelets at 15 DAA. Grains were ground with liquid nitrogen and 40mg of pulverized grains were added to 100ul of SD-PAGE loading protein buffer 2X at 95°C for 10 min and centrifuged at 13000 g for 10 min for protein extraction. 5ul aliquots were loaded per lane in a 12% SDS-PAGE. Gels were stained with SimpleBlue SafeStain (safe Coomassie G-250). In-gel tryptic digestion was performed, as previously described (De Pablos *et al.*, 2019). Peptides were loaded onto an UltiMate 3000 RSLCnano HPLC system (Thermo) equipped with a PepMap 100 Å C18, 5 µm trap column (300 µm x 5 mm Thermo) and a PepMap, 2 µm, 100 Å, C18 EasyNano nanocapillary column (75 µm x 150 mm, Thermo) and separated as previously described (De Pablos *et al.,* 2019), with the exception that the following gradient was used: 3-10% B over 8 mins, 10-35% B over 125 mins, 35-65% B over 50 mins, 65-99% B over 7 mins and then proceeded to wash with 99% solvent B for 4 min. The nanoLC system was interfaced with an Orbitrap Fusion hybrid mass spectrometer (Thermo) with an EasyNano ionisation source (Thermo). Positive ESI-MS and MS^2 spectra were acquired using Xcalibur software (version 4.0, Thermo) as previously described (De Pablos *et al.*, 2019). Tandem mass spectra were searched against the *Triticum aestivum* subset of the UniProt database appended with the sequence of *TaEXPA6* using Mascot program (Matrix Science Ltd., version 2.6.1). Search criteria specified: Missed cleavages, 1; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M); Peptide tolerance, 3 ppm; MS/MS tolerance, 0.5 Da. Results were filtered to 2% FDR against a reversed database and required a minimum of two unique peptides per protein. Label-free peptide quantification was extracted from aligned precursor ion areas using Progenesis QI (Version 2.2., Waters).

Statistical analysis of data

ANOVA was applied when multiple groups of data were compared followed by pairwise comparisons between the control and each T line (Fisher`s LSD test post-hoc) to evaluated significant differences ($p<0.10$, $p<0.05$, $p<0.01$ and $p<0.001$) using the STATISTICA 7 software. Regression analyses were used to evaluate the degree of association between variables.

Results

We generated transgenic wheat plants expressing *TaExpA6*, an α-expansin normally expressed in wheat roots (according to wheat-expression browser: Ramírez-González *et al.*, 2018), under the control of the wheat puroindoline-b (PinB) gene promoter (*TraesCS7B02G431200*), which drives transcription restricted to the endosperm, aleurone and pericarp layers in developing seeds, but not in the embryo (Gautier *et al.*, 1994; Digeon *et al.*, 1999). We confirmed that the expression of the PinB gene is negligible outside the mentioned tissues using the wheat expression browser (Ramírez-González, *et al*., 2018). We further confirmed this by RT-qPCR and no expression of the transgene was found in roots and vegetative tissues (data not shown). Homozygous T2 lines were selected and T3 and T4 generations were evaluated in field experiments at different plant densities. Low density (44 plants m^{-2}) plantings were carried out during the 2016-17 growing season using T3 seed; while a more typical agronomic density (300 plants m^{-2}) experiment was carried out using T4 seed in the 2018-19 growing season in Valdivia (Chile). Transgene expression in developing seeds was assessed by RT-qPCR and the levels of recombinant *TaExpA6* protein were assessed by semi-quantitative proteomic analysis.

TaExpA6 expression and protein abundance in growing seeds of T3 and T4 transgenic lines and segregating wild type plants

Twenty-four transgenic lines were generated and evaluated in a field experiment together with wild type of cultivar Fielder that had been through the same tissue culture cycle as the transformants (control) at low plant density (44 plants m^{-2}), aimed also at bulking seed for further studies. Based on the transcript abundance and transgene copy number (Tables S1,

S2 and Fig. S2), four transformed lines were chosen for further analysis. Transgene transcripts of *TaExpA6* were not detected in growing seeds at 5 days after anthesis (DAA) and showed a peak in abundance between 10 and 20 DAA (Fig. 2). In both experiments, all the transgenic lines showed *TaExpA6* transcripts, which were not apparent in the control (Fig. 3a). Transgenic lines 4 and 3 showed the highest relative transcript abundance (Fig. 3a).

The presence of *TaExpA6* protein was assessed by extracting total proteins from developing seeds at 15 DAA (see Methods). Lizana *et al*. (2010) showed that several α-expansin genes are normally expressed in developing seeds and distinguishing between these closely related proteins is difficult using immunological methods. Instead, we used LC-MS/MS proteomic methods, which allowed unambiguous protein identification as well as the determination of relative protein abundance. The abundance of the recombinant *TaExpA6* protein showed good agreement with the observed transgene expression levels in transgenic lines, with Lines 3 and 4 showing the highest recombinant protein abundance (Fig. 3b).

TaExpA6 ectopic expression increases grain size and weight in transgenic wheat

In our growth environment each central spikelet of one ear of wheat typically accommodates four grains, named G1-G4 based on the relative proximity to the rachis of the spike. The more distal grains are typically the smallest with G1 and G2 typically the biggest. The examination of individual grain weight from our transgenic lines revealed significant increases in all four grain positions of Line 4 in the low plant density experiment (Fig. 4a). Interestingly, the biggest GW increases were seen in the more distal position (G4), which showed increases of up to 32.8 % in Line 4 compared to control, whereas in G1 the increase was 12.3 % with similar values in G2 (9.5 %) and G3 (12.2 %). Similar but non-significant increases in GW were seen in the other transgenic lines.

The high plant density field experiment was performed using T4 seed planted at 300 seed m^{-2}, which is generally used by farmers for spring wheat in southern Chile. In this separate generation and independent field experiment, similar increases in individual GW were evident in the transgenic lines compared to the control. Transgenic Line 4 again showed the

best performance in this experiment, outperforming the control in terms of GW by 11.6, 10.9, 11.3 and 26.6 % in G1, G2, G3 and G4, respectively, with the biggest increases seen in G4 (Fig. 4b). The other transgenic lines showed increased GW compared to control in the distal grain positions G4 and at G1, G2 and G3 depending on the line and experiment (Fig. 4a and b).

Higher GW is likely associated with increased size, and to examine this we compared the length, width and surface area of grains from transgenic and control lines from the high density experiment. The results of grain dimensions of G2 (Fig. 5) showed that increased GW was most closely associated with increased grain length and area with little increase in grain width, and similar results were found in the other grain positions (Fig. S3). When the associations between GW and grain length, width or area were plotted across genotypes and grain positions, higher consistency and better residual distributions were found for the relationship between GW and grain length (Fig. 6). This indicates that increased grain length is the largest contributor to higher grain area in transgenic lines.

TaExpA6 ectopic expression increases average grain weight and grain yield in transgenic wheat

The ectopic expression of*TaExpA6* increased average GW in the transgenic lines in both experiments, and this effect translated into increased total GY. Once again, the biggest increase in total yield was apparent for transgenic Line 4, which showed the largest transcript and protein abundance of *TaExpA*6 during early grain development. Grain yield per plant for Line 4 was 9.5 % higher than that of control, although this increase was not statistically significant ($P > 0.10$) in the low density experiment (Table 1). However, GY was 11.3 % higher than the control per square meter in the higher density field experiment (Table 2). In the higher plant density experiment, we monitored GY, GN and average GW for both stem categories (main stems and tillers). We observed no significant differences in total GN ($P>0.10$) between control and transgenic lines, and the observed increases in overall yield (11.3 %) that are remarkably similar to those of individual GW in these experiments. These data reveal that there is no evident trade-off between the major yield components (weight and number) in our experiments as shown in Figure 7. Notably,

developmental phases and stages such as dates of anthesis and physiological maturity were similar among the transgenic lines and control in all experiments (Table S3). Line 4 was slightly taller ($p<0.01$) than the control in low and regular density experiments (Table S4). Line 1 also showed increased plant height, but only in the low density experiment, whereas Lines 2 and 3 did not show differences in plant height compared to control across experiments (Table S4).

Discussion

The trade-off between GW and GN has been reported in many studies (e.g. Foulkes *et al*., 2011; Molero *et al.,* 2019 and references therein). From a physiological point of view, several studies have suggested that the negative correlation between GN and GW derives from the large proportion of "small grains" at distal positions due to either wheat breeding or crop management, and is independent of any competitive relationship among developing grains (Acreche and Slafer 2006; Ferrante et al. 2015, 2017). This hypothesis is supported by the low correlation reported between final GW and starch-synthesizing enzymes (Fahy et al. 2018). However, this does not explain why the successful attempts to increase GW have been accompanied by compensatory decreases in GN (e.g. Wiersma *et al.*, 2001; Brinton *et al.*, 2017; Wang *et al.*, 2018). This negative relationship between weight and number was recently confirmed as a general phenomenon in analyses across a range of elite wheat genotypes (Quintero *et al*., 2018; Molero *et al*., 2019). A promising candidate gene was recently identified in wheat underlying a QTL that controls spikelet number per spike (Kuzay *et al.*, 2019). Near isogenic lines carrying this gene increased GN, but GW concomitantly decreased by 19 %, preventing an increase in yield. Few studies have made efforts to understand the molecular and genetic basis of the trade-off between both main yield components in wheat. In tetraploid wheat, Golan et al. (2019) have suggested that gsn1 (grain size and number 1), FZP (FRIZZY PANICLE) and GNI-1 (Grain Number Increase 1) are involved in coordinating the trade-off between GN and GW by integrating cell differentiation and proliferation processes. In addition, trehalose6-phosphate/SnRK1 has been suggested to influence grain yield by increasing the potential rate of filling and grain mass but neither GN nor yield were reported by the authors (Zhang et al., 2017). Therefore, the key genes, moment, and cellular location of this coordination remain largely

undefined. This background suggests that it may be difficult to disrupt the complex regulatory pathways that control these crop traits. However, our results demonstrate that increasing the level of α-expansin protein during early grain development leads to increased grain size in wheat as a result of increased grain length.

In our experiments, bigger grains resulted in increased total yield as there is no associated compensatory decrease in GN. This contrasts with previous attempts to increase grain size using conventional wheat breeding, QTL or mutants (Wiersma *et al*., 2001; Brinton *et al*., 2017; Wang *et al*., 2018; Molero *et al*., 2019). Bae *et al.*, 2014 reported increases in seed weight in Arabidopsis by the expression of a sweet potato expansin. In this case, expansin expression was driven by the broadly expressed cauliflower mosaic virus *35S* promoter and led to widespread morphological changes in the number and size of leaves, number of siliques, etc. in the plants. Since Arabidopsis is essentially a wild weedy plant, with no previous breeding for yield components, it is hard to draw any conclusions on agronomically relevant yield from those experiments. The use of a specific grain promoter in our work confined the expression of the *PinB::TaExpA6* transgene to the developing grain without detrimental pleiotropic impact on plant growth and development that might impair crop performance.

Our study reveals that it is possible to break the negative association between GW and GN using a targeted transgenic approach, at least in the high yielding environment of southern Chile. For many years GN and GW were assumed to be independent of each other. However, it has been demonstrated that the developmental determination of these two key yield components shows a close temporal overlap in wheat, which occurs between booting and 10 days after anthesis (Calderini *et al.*, 1999; Ugarte *et al.*, 2007; Brinton & Uauy, 2019). These findings were confirmed by a recent study of wheat cultivars across different environments in Australia (Parent *et al.*, 2017). The overlap between the determination of GW and GN is similarly apparent in other grain crops such as barley, triticale, sorghum and sunflower (Lindström *et al.*, 2006; Ugarte *et al.*, 2007; Yang *et al.*, 2009; Castillo *et al.*, 2017). The temporal overlap between GN and GW determination suggests that they are

developmentally linked, giving rise to the observed trade-off between the two yield components before grain filling begins.

Our targeted approach, using an early grain promoter to drive ectopic expression of an expansin gene in young developing grain, may have been successful due to the timing of the *PinB::TaExpA6* expression. The lack of expression of the *PinB::TaExpA6* at 5 DAA and the peak levels observed between 10 and 20 DAA in our study, suggests that the fruitful increase of GW without a negative impact on GN was due to the expression of the expansin transgene occurring after the overlapping period of GW and GN determination, thereby avoiding the trade-off between these yield components (see Fig. 1). However, only two of the four lines showed significant positive impact on grain yield (Lines 1 and 4) despite all four lines showing *TaExpA6* expression in the grain. The most likely explanation for this is that a threshold amount of additional expansin is needed in order to see a significant effect, as the lines with highest *TaExpA6* protein in developing grain, showed the most significant increases in yield. It is, however, also possible that some of the differences between the lines may be the result of different transgene integration sites, which may impact on the expression of other genes in the region.

In our experiments GY was increased in both low and high density plantings, but the effect was greater at higher plant density. In the higher density experiment, we assessed the impact on grain in both the main stem and tillers and observed that the effects of the transgene were smaller in grains from tillers, i.e. for Line 4, GY of main stems increased almost 19 % ($P < 0.05$) and tillers only 5 % ($P > 0.10$). This may explain why GY was higher in the high density experiment than under low planting densities, where more tillers are typically produced.

Our results are a very encouraging demonstration that the trade-off between GW and GN can be broken and, as a consequence, GY can be increased. However, we recognize that more experiments across different environments and cultivars should be carried out to confirm the results presented here, which should be seen as a proof of concept of the positive effect of ectopic expansin expression in developing wheat seeds.

This work provides a simple approach for breaking barriers in wheat yield that may also prove important in a wide range of crops where the trade-off between GN and grain size is widely observed.

Acknowledgements

We would like to thank Andy Greenland for his encouragement and advice and Gustavo Slafer (ICREA, University of Lleida) for comments on an early version of Figure 1. We also thank Victoria Abarzua, Gabriela Carrasco, Magda Lobnik and the personnel of the experimental field station of the Universidad Austral de Chile (EEAA) for their technical assistance. Wheat transformation was funded by BBSRC grant BB/J019356/ to NIAB. Experiments in Chile were partially funded by the Sustainable Modernization of Traditional Agriculture (MasAgro) Project initiative from the Secretariat of Agriculture and Rural Development (SADER) and CIMMYT (Mexico) and the Chilean Technical and Scientific Research Council (CONICYT) by FONDECYT Project 1170913. Work at the University of York was supported by BBSRC grant. MC, SB and MM were funded by BBSRC grant BB/J019356/ awarded to Andy Greenland and EW. International Stay Fellowship of Francisca M. Castillo at The University of York (2015) was funded by CONICYT.

Author contributions

D.F. Calderini, conceived the project, coordinated the field trials, wheat transformation, experiments and data analyses; F.M. Castillo, designed experiments, genetic constructs and analysed the experimental data; A. Arenas-M, evaluated expansin expression and molecular data analyses; G. Molero and M.P. Reynolds collaborated in designing the low plant density experiment and data analysis; M. Craze, S. Bowden and M. Milner carried out the transformation of wheat; E. Wallington, supervised wheat transformation; A. Dowle, performed the proteomic and data analysis for determining expansin abundance; L.D. Gomez, coordinated the experiments, evaluated transformants, designed experiments, genetic constructs and data analyses; S.J. McQueen-Mason, conceived and coordinated the project, wheat transformation, experiments and data analysis. D.F. Calderini, F.M. Castillo, L.D. Gomez and S.J. McQueen-Mason wrote the manuscript with contributions from all authors.

References

Acreche MM, Slafer GA. **2006**. Grain weight response to increases in number of grains in wheat in a Mediterranean area. *Field Crops Research* **98**: 52–59.

Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, Pozniak CJ, Choulet F, Distelfeld A, Poland J, *et al.* 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361 (6403) eaar7191**.

Bae JM, Kwak MS, Noh SA, Oh MJ, Kim YS, Shin JS. **2014**. Overexpression of sweetpotato expansin cDNA (IbEXP1) increases seed yield in Arabidopsis. *Transgenic Research* **23**: 657–667.

Borrás L, Slafer GA, Otegui ME. **2004**. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: A quantitative reappraisal. *Field Crops Research* **86**: 131–146.

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 Phytologist* **215**: 1026–1038.

Brinton J, Uauy C. **2019**. A reductionist approach to dissecting grain weight and yield in wheat. *Journal of Integrative Plant Biology* **61**: 337–358.

Bustos D V., 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 Research* **145**: 106-115.

Calderini DF, Abeledo LG, Savin R, Slafer GA. **1999**. Effect of temperature and carpel size during pre-anthesis on potential grain weight in wheat. *Journal of Agricultural Science* **132**: 453–459.

Calderini DF, Savin R, Abeledo LG, Reynolds MP, Slafer GA. **2001**. The importance of the period immediately preceding anthesis for grain weight determination in wheat. *Euphytica* **119**: 199–204.

Castillo FM, Vásquez SC, Calderini DF. **2017**. Does the pre-flowering period determine the potential grain weight of sunflower? *Field Crops Research* **212**: 23–33.

Choi D, Lee Y, Cho HT, Kende H. **2003**. Regulation of expansin gene expression affects growth and development in transgenic rice plants. *Plant Cell* **15**: 1386–1398.

Cosgrove DJ. **2018**. Nanoscale structure, mechanics and growth of epidermal cell walls. *Current Opinion in Plant Biology* **46**: 77–86.

Digeon JF, Guiderdoni E, Alary R, Michaux-Ferrière N, Joudrier P, Gautier MF.

1999. Cloning of a wheat puroindoline gene promoter by IPCR and analysis of promoter regions required for tissue-specific expression in transgenic rice seeds. *Plant Molecular Biology* **39**: 1101–1112.

Fleming AJ, McQueen-Mason S, Mandel T, Kuhlemeier C. **1997**. Induction of leaf primordia by the cell wall protein expansin. *Science* **276**: 1415–1418.

Foulkes MJ, Slafer GA, Davies WJ, Berry PM, Sylvester-bradley R, Martre P, Calderini DF, Griffiths S, Reynolds MP. **2011**. Raising yield potential of wheat . III . Optimizing partitioning to grain while maintaining lodging resistance. **62**: 469–486.

García GA, Serrago RA, González FG, Slafer GA, Reynolds MP, Miralles DJ. **2014**. Wheat grain number: Identification of favourable physiological traits in an elite doubled-haploid population. *Field Crops Research* **168**: 126–134.

Gautier MF, Aleman ME, Guirao A, Marion D, Joudrier P. **1994**. Triticum aestivum puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant Molecular Biology* **25**: 43–57.

Griffiths S, Wingen L, Pietragalla J, Garcia G, Hasan A, Miralles D, Calderini DF, Ankleshwaria JB, Waite ML, Simmonds J, *et al.* 2015. Genetic dissection of grain size and grain number trade-offs in CIMMYT wheat germplasm. *PLoS ONE* **10**: 1–18.

Gross JD, Moerke NJ, Von Der Haar T, Lugovskoy AA, Sachs AB, McCarthy JEG, Wagner G. **2003**. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115**: 739–750.

Herrera J, Calderini DF. 2020. Pericarp growth dynamics associate with final grain weight in wheat under contrasting plant densities and increased night temperature. *Annals of Botany.* In press. doi.org/10.1093/aob/mcaa131.

Hood EE, Gelvin SB, Melchers LS, Hoekema A. 1993. New Agrobacterium helper plasmids for gene transfer to plants. *Transgenic Research*. **2**, 208–218.

Howells RM, Craze M, Bowden S & Wallington EJ. 2018. Efficient generation of stable, heritable gene edits in wheat using CRISPR/Cas9. *BMC plant biology*, **18**, 215.

Ishida Y, Tsunashima M, Hiei Y & Komari T. 2014. Wheat (*Triticum aestivum* L.) Transformation Using Immature Embryos. *Agrobacterium protocols*. **1223**, 189-198.

Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N, *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. *Theoretical and Applied Genetics* **132**: 2689–2705.

Lindström LI, Pellegrini CN, Aguirrezábal LAN, Hernández LF. **2006**. Growth and

development of sunflower fruits under shade during pre and early post-anthesis period. *Field Crops Research* **96**: 151–159.

Lizana XC, Riegel R, Gomez LD, Herrera J, Isla A, McQueen-Mason SJ, Calderini DF. **2010**. Expansins expression is associated with grain size dynamics in wheat (Triticum aestivum L.). *Journal of Experimental Botany* **61**: 1147–1157.

Mcqueen-Mason S, Cosgrove DJ. **1994**. Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 6574–6578.

McQueen-Mason S, Durachko DM, Cosgrove DJ. **1992**. Two endogenous proteins that induce cell wall extension in plants. *Plant Cell* **4**: 1425–1433.

Molero G, Joynson R, Pinera-Chavez FJ, Gardiner LJ, Rivera-Amado C, Hall A, Reynolds MP. **2019**. Elucidating the genetic basis of biomass accumulation and radiation use efficiency in spring wheat and its role in yield potential. *Plant Biotechnology Journal* **17**: 1276–1288.

De Pablos LM, Ferreira TR, Dowle AA, Forrester S, Parry E, Newling K, Walrad PB. **2019**. The mRNA-bound Proteome of Leishmania mexicana: Novel genetic insight into an ancient parasite*. *Molecular and Cellular Proteomics* **18**: 1271–1284.

Parent B, Bonneau J, Maphosa L, Kovalchuk A, Langridge P, Fleury D. **2017**. Quantifying wheat sensitivities to environmental constraints to dissect genotype × environment interactions in the field. *Plant Physiology* **174**: 1669–1682.

Pien S, Wyrzykowska J, McQueen-Mason S, Smart C, Fleming A. **2001**. Local expression of expansin induces the entire process of leaf development and modifies leaf shape. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 11812–11817.

Quintero A, Molero G, Reynolds MP, Calderini DF. **2018**. Trade-off between grain weight and grain number in wheat depends on GxE interaction: A case study of an elite CIMMYT panel (CIMCOG). *European Journal of Agronomy* **92**: 17–29.

Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A, *et al.* 2018. The transcriptional landscape of polyploid wheat. *Science* **361**.

Reynolds M, Foulkes MJ, Slafer GA, Berry P, Parry MAJ, Snape JW, Angus WJ. **2009**. Raising yield potential in wheat. *Journal of Experimental Botany* **60**: 1899–1918.

Risacher T, Craze M, Bowden S, Paul W, Barsby T. 2009. Highly efficient Agrobacterium-mediated transformation of wheat via in planta inoculation. *In Transgenic*

Wheat, Barley and Oats: Production and Characterization Protocols, Methods in Molecular Biology, Humana Press; Totowa, NJ, USA **478**: 115-124.

Rivera-Amado C, Trujillo-Negrellos E, Molero G, Reynolds MP, Sylvester-Bradley R, Foulkes MJ. **2019**. Optimizing dry-matter partitioning for increased spike growth, grain number and harvest index in spring wheat. *Field Crops Research* **240**: 154–167.

Rochange SF, Wenzel CL, Mcqueen-Mason SJ. **2001**. Impaired growth in transgenic plants over-expressing an expansin isoform. *Plant Molecular Biology* **46**: 581–589.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, Hoff MJB Van Den, Moorman AFM. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research.* **37:** 1-12.

Sangha JS, Gu K, Kaur J, Yin Z. **2010**. An improved method for RNA isolation and cDNA library construction from immature seeds of *Jatropha curcas* L. *BMC Research Notes*, **3:** 2–7.

Simmonds J, Scott P, Leverington-Waite M, Turner AS, Brinton J, Korzun V, Snape J, Uauy C. **2014**. Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biology* **14**: 1–13.

Slafer GA, Savin R. 1994. Source-sink relationships and grain mass at different positions within the spike in wheat. *Field Crops Research* **37**: 39–49.

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. **2007**. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics* **39**: 623–630.

Ugarte C, Calderini DF, Slafer GA. **2007**. Grain weight and grain number responsiveness to pre-anthesis temperature in wheat, barley and triticale. *Field Crops Research* **100**: 240–248.

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. *Theoretical and Applied Genetics* **131**: 2463–2475.

Wiersma JJ, Busch RH, Fulcher GG, Hareland GA. **2001**. Recurrent selection for kernel weight in spring wheat. *Crop Science* **41**: 999–1005.

Yang Z, Van Oosterom EJ, Jordan DR, Hammer GL. **2009**. Pre-anthesis ovary development determines genotypic differences in potential kernel weight in sorghum. *Journal of Experimental Botany* **60**: 1399–1408.

Zadoks JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. *Weed research.* **14(6)**:415-21.

Zhang P, He Z, Tian X, Gao F, Xu D, Liu J, Wen W, Fu L, Li G, Sui X, et al. 2017. Cloning of TaTPP-6AL1 associated with grain weight in bread wheat and development of functional marker. *Molecular Breeding* **37**: 78, 1–8.

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

Figure Legends

Figure 1. Schematic description of processes and traits of grain weight determination in wheat from booting to maturity. The gene expression of the recombinant *PinB::TaExpA6* and its apparent dynamic is shown. At the bottom of the scheme the overlap between grain number and grain weight determination from booting to the end of the lag phase are shown. Wide arrows show the main links between processes/traits and narrow arrows indicate indirect links between processes/traits.

Figure 2. Relative expression of *TaExpA6* in grains at 5, 10, 15, 20 and 25 days after anthesis (dpa) in the field experiment at regular agronomical plant density of 300 pl m^{-2} assessed by RT-qPCR. Control line correspond to spring wheat cultivar 'Fielder' that have undergone the same tissue culture process as the transformed lines.
Asterisks indicate significant differences by pairwise comparisons between each line and the Control (Fisher`s LSD test post hoc). Significant effects at $p<0.10$ (*); $p<0.05$ (**); $p<0.01$ (***); $p<0.001$ (****); ns: not significant. All data are shown as mean and standard error.

Figure 3. Expression and protein abundance of *TaExpA6* in grains. Control line correspond to spring wheat cultivar 'Fielder' that underwent the same tissue culture process as the transformed lines. **(a)** Relative mRNA levels of the *TaExpA6* transgene assessed by RT-qPCR in grains at 15 days after anthesis (DAA) in the Control and transformed lines 1, 2, 3 and 4 from experiment with low (LD) and regular plant density (RD). **(b)** Relative protein abundance as assessed by LC-MS/MS analysis at 15 DAA in the Control and transgenic lines 1, 2, 3 and 4 at regular agronomical plant density. Asterisks indicate significant differences by pairwise comparisons between each line and the Control (Fisher`s LSD test post hoc). Significant effects at $p<0.10$ (*); $p<0.05$ (**); $p<0.01$ (***); $p<0.001$ (****); ns: not significant. All data are shown as mean and standard error.

Figure 4. Individual grain weight. Grain weight at grain position 1 (G1), 2 (G2), 3 (G3) and 4 (G4) in the Control and *TaExpA6* transgenic Lines 1, 2, 3 and 4 at **(a)** low plant density and **(b)** regular plant density. The control line corresponds to spring wheat cultivar 'Fielder' that

underwent the same tissue culture process as the transformed lines. Asterisks indicate significant differences evaluated by pairwise comparisons between each line and the Control (Fisher`s LSD test post hoc). Significant effects at $p<0.10$ (*); $p<0.05$ (**); $p<0.01$ (***); $p<0.001$ (****); ns: not significant. All data are shown as mean and standard error.

Figure 5. Grain length, width and area at grain position 2 (G2) in wild type and transformed wheat lines. Grain dimensions were evaluated in control and four *TaExpA6* transgenic lines (lines 1 – 4) in field experiments at low plant density of 44 pl m^{-2} and regular agronomical plant density of 300 pl m^{-2}. **(a)** Grain length, **(b)** grain width and **(c)** grain area. The control line corresponds to spring wheat cultivar 'Fielder' that underwent the same tissue culture process as the transformed lines. Asterisks indicate significant differences evaluated by pairwise comparisons between the control and each transgenic line (Fisher`s LSD test post hoc). Significant effects at $p<0.10$ (*); $p<0.05$ (**); $p<0.01$ (***); $p<0.001$ (****); ns: not significant. All data are shown as mean and standard error.

Figure 6. Association between individual grain weight and dimensions. Grain weight and grain length **(a)**, width **(b)** and area **(c)** of grain positions 1 (G1: circles), 2 (G2: square), 3 (G3: rhombus) and 4 (G4: triangles) from central spikelets of the spike in the control (black) and transgenic lines 1 (white), 2 (green), 3 (orange) and 4 (red) recorded from the field experiment at plant density of 300 pl m^{-2}.

Figure 7. Trade-off between grain weight and grain number. Relationship between grain weight and grain number of the control line (black circles) and transgenic lines 1 (open white circles), 2 (green circles), 3 (orange circles) and 4 (red circles) in the low plant **(a)** and regular agronomical plant **(b)** density experiments. The regression line (continuous black line) and two confidence bands surrounding the best-fit line that define the confidence interval (dotted lines) are shown.

Legends to Supplementary material

Fig. S1 Schematic diagram of the binary plasmid pEW279-Exp T-DNA.

Fig. S2. Screening of relative expression of transgene in 15 wheat transgenic lines at T2 generation.

Fig. S3. Box and whiskers showing a) grain weight, b) grain length, c) grain width and d) grain area across grain positions (G1-G4) in control line and transgenic lines (1-4).

Table S1. Selection criteria of four wheat transgenic lines to perform experiments at low and regular plant density

Table S2. Transgene copy number determined by NPTII amplification.

1 **Table S3.** Phenology of transgenic wheat lines and control in both experiments.

2 **Table S4.** Plant height of transgenic wheat lines and control in both experiments.