Characterization of Potential Candidate Genes for Grain Size in Wild Emmer Wheat *Triticum Dicoccoides*

Sanket Shinde
Punjab Agricultural University
Guriqbal Singh Dhillon
Punjab Agricultural University
Amandeep Kaur
Punjab Agricultural University
Parveen Chhuneja
Punjab Agricultural University
Achla Sharma
Punjab Agricultural University
Satinder Kaur (satinder.biotech@pau.edu)
Punjab Agricultural University https://orcid.org/0000-0003-3704-3074

Research Article

Keywords: Triticum dicoccoides, Emmer wheat, Grain size, Grain length, Expression studies, wheat improvement

Posted Date: July 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-705356/v1

Abstract

There is an incessant need to address food security in staple crops, and the crop yield is positively correlated with grain weight. Grain size, determined by grain length and width, is an essential component of final grain weight in cereals. Wheat wild relatives are the goldmine to harness any trait of interest, including the component traits of grain size. It is crucial to understand the detailed mechanism of grain size formation and unravel underlying genes controlling grain size in these species for their proper utilization in wheat improvement. In this study, gene expression analysis was performed on developing grain in wild tetraploid progenitor *Triticum dicoccoides* (AABB) to identify candidate genes involved in determining grain size. Four *T. dicoccoides* accessions were selected, two (pau5228 and pau5322) with higher grain length and weight and two (pau14703 and pau14756) with comparatively smaller grains.

Six genes out of the eight genes selected for expression study, viz., *GL7*, *TaGL3*, *TaGS5*, *GS3*, *SRS3*, and *TaGASR7,* were upregulated from 8 days post-anthesis (DPA) to 20 DPA in both the large grain accessions, while *TaGW2* gene was upregulated in both small grain accessions. *TGW6* was downregulated in all the accessions at all stages of grain development. The results indicated that the selected genes play an essential role in grain size formation by controlling individual morphometric components of grain length and width. Targeted introgression genes controlling grain size components will eventually aid in improving grains yield.

Declarations

Ethics approval: NA

Consent to participate: NA

Consent for publication: All the authors have agreed to the publication of this manuscript

Availability of data and material: NA

Code availability: NA

Author contribution statement

SK, AS and PC conceived and designed the research. SS conducted the experiments and collected the data. SS, GSD, and AK produced the final draft of the manuscript. SK, AS, and PC reviewed the experiments and manuscript.

Acknowledgments NA

Conflicts of interest The authors declare that they have no conflict of interest.

Introduction

Wheat is one of the major global cereal grains that occupies seventeen percent of the crop area worldwide, feeding approximately 30% of the world's population (Eversole et al. 2014). There is a hike in demand for food production as expected by the continual expansion of the population, which would be 9 billion by 2050, thus increasing pressure on agricultural systems (FAO et al. 2015). Increasing crop yield is the only sustainable route towards meeting this demand. However, a decline in rates of yield has been observed, which is a major bottleneck to accomplish the expected doubling in crop production (Ray et al. 2013).

Grain yield in wheat is a complex trait, controlled by several components like productive spike number per unit area, grain number per spike, grain size, and thousand grain weight (TGW) (Zhang et al. 2015). TGW in wheat is the most stable yield component and is positively correlated with crop yield, highlighting the need to carry out studies regarding genetic control to enhance breeding efficiency (Kuchel et al. 2007; Yu et al. 2019). TGW is mainly determined by individual grains size and morpho-metric components of grain length, width, and area. During the domestication process, the grain size was the primary target that has been widely selected and manipulated to increase grain yield (Gegas et al. 2010).

In the last decade, functional genomics has increased the understanding of grain size. A number of genes associated with grain size/weight have been identified in model crops like rice and *Arabidopsis* (Zuo and Li 2014). Several signaling pathways and diverse mechanisms influencing the endosperm and maternal tissue growth determining the seed size have been reported (Li and Li 2016). A number of genes are identified in major metabolic pathways and are involved in cell expansion and cell division regulation. These genes might be functional in the ubiquitination-mediated proteasomal degradation pathway (*GW2* and *GW5/QsW5*), G-protein signaling (*GS3* and *DEP1*), phytohormones (*TGW6, CKX2,* and *GS6*), and/or other unknown pathways (*GS5* and *GW8*) (Huang et al. 2009; Li et al. 2011; Mao et al. 2010; Shomura et al. 2008; Song et al. 2007; Wang et al. 2012). Several genes in rice have been proven to influence grain size and weight, such as *GS5* (Li et al. 2011; Xu et al. 2015), *TGW6* (Ishimaru et al. 2013), *GASR7* (Li and Yang 2017), *qGL3* (Qi et al. 2012; Zhang et al. 2012), *GS3* (Fan et al. 2006), *GW2* (Song et al. 2007), *GW5* (Weng et al. 2008), *GL7* (Wang et al. 2015b), *GW7* (Wang et al. 2015a), *GLW7* (Si et al. 2016), *GS2* (Duan et al. 2015; Hu et al. 2015), *SLG7* (Zhou et al. 2015), *SRS5* (Segami et al. 2012) and *SRS3* (Kitagawa et al. 2010).

Despite these advances, an understanding of the control of grain size is limited in wheat. However, some QTLs for grain weight and size have also been characterized in wheat (Breseghello and Sorrells 2007; Brinton et al. 2017; Gegas et al. 2010; Kumar et al. 2016; Simmonds et al. 2014), a few of them have been cloned. Several orthologs of rice have been characterized in wheat controlling grain size-related traits such as *TaGW2* (Bednarek et al. 2012; Simmonds et al. 2016; Su et al. 2011; Yang et al. 2012), *TaGASR7* (Zhang et al. 2014), *TaGS5* (Ma et al. 2016), *TaTGW6* (Hanif et al. 2016) and *TaGL3* (Yang et al. 2019). Traits like grain length, grain width, and grain thickness are often positively correlated with grain size, and these traits are further used to evaluate the grain weight in the breeding programs. Also, grain weight has been more extensively studied in wheat than subcomponents like grain length and width. Hence, there is a need to look into the genetic mechanism controlling these yield components, ultimately contributing to overall yield. Studying the grain development process seems essential to find the genetic architecture and genes involved in grain size formation. Understanding the regulatory mechanisms underlying early gene expression and

selecting the candidate genes for grain size is of great significance for yield and quality improvement in wheat (Guan et al. 2019).

There has been a radical change in wheat genomics in the last few years in terms of resources such as transcriptomic databases (Borill et al. 2016; Pearce et al. 2015), availability of an annotated wheat genome (IWGSC 2018), high-quality gene models (Clavijo et al. 2017) and high-density single nucleotide polymorphism (SNP) arrays (Wang et al. 2014, Winfield et al. 2016). Proteomic and transcriptomic studies have provided a global overview of genes involved in grain development (Brinton et al. 2018; Wan et al. 2008). Comparative studies with rice also paved the way to clone and characterize a significant number of genes in wheat (Brinton and Uauy 2019). Identifying such potential genes can lead to the establishment of novel combinations of distinct mechanisms and variations across homeologs that influence grain size.

Significant variations in grain size and weight occur among wild species of diploid, tetraploid, and hexaploid wheat (Gegas et al. 2010). *T. dicoccoides* is known as wild emmer wheat, the progenitor of cultivated wheat, contributing to grain size during wheat evolution (Feldman and Kislev 2007). Punjab Agricultural University has a collection of more than 110 *T. dicoccoides* accessions, and a study on its grain size traits from previous years indicated that this germplasm is a good source of variation for grain size-related traits. The present investigation aimed to understand the role of some potential genes responsible for controlling grain length and grain width in different accessions of *T. dicoccoides* having variations in grain length and width.

Materials And Methods

Plant material

Plant material used includes four *T. dicoccoides* accessions, selected on the basis of contrasting thousand grain weight (TGW). Two of these accessions, pau5228 and pau5232, originated in Turkey, had larger grains and hereon called LG accessions. The other two accessions, pau14703 and pau14756, with the origin in Israel, had smaller grains hereon called SG accessions. The four selected accessions were evaluated for grain size parameters in developing grain planted in three replications, in completely randomized design (CRD) for two consecutive years, 2018-19 and 2019-20 at experimental field area, School of Agricultural Biotechnology, PAU, Ludhiana.

Measurement of grain size

The grain size was determined both in the developing and mature grains. The developing grains were collected at seven different stages (post-anthesis); the first sample is collected at 4-days post-anthesis (4 DPA), then at 8DPA, 12DPA, 16DPA, 20DPA, 24DPA, 28DPA, and mature grains for two consecutive wheat seasons 2018-19 and 2019-20. The spikes in each accession of *T. dicoccoides* were tagged with the anthesis date to maintain uniformity in the sampling marking. Developing seeds were collected only from the primary florets of three spikelets selected from the middle of each spike. In the developing grain, grain length (GL), grain width (GW), and grain area (GA) of collected seeds were measured, while in mature grain GL, GW and GA were taken using the Canon5600 scanner GrainScan software (Whan et al. 2014). As

manual threshing of one thousand grains in wild accessions of *T. dicoccoides* spikelets is difficult due to hard threshing nature, hundred-grain weight was recorded and converted into thousand-grain weight (TGW). The data was analyzed in the statistical software R (R Core Team 2018), using the "AOV" function for variance and Duncan Multiple Range Test (DMRT) to compare mean values of four *T. dicoccoides* accessions. For further analysis, the adjusted means of two environments were represented as the third environment. The graphical representation of the data was plotted using "GGplot2" version 3.3.2 and "GGpubr" version 0.4.0 packages of Rstudio (Kassambara and Kassambara 2020; Wickham 2016).

Identification of candidate gene for grain size

Candidate genes for grain size component traits were selected from available literature in wheat and rice. Nucleotide sequences of complete genes were retrieved from respective RGAP (Rice Genome Annotation Project) and NCBI (National Centre for Biotechnology Information) database and were used as a query for online BLAST against RefSeqv1.0 of wheat. Sequences showing maximum alignment with A and B genomes chromosomes were selected, and again BLAST was done against *T. dicoccoides* genome in Ensembl Plants database (http://www.ensemblgenomes.org/id/%s). The final sequences obtained were filtered based on low e-value, query coverage, and gene functional annotation.

RNA extraction and cDNA synthesis

Total RNA was isolated from developing seeds using RNAiso Plus reagent (Takara, Japan) in three technical replicates. The concentration of RNA was measured by spectrophotometry using Nanodrop™ 1000 (Thermo Scientific, USA), and quality was confirmed with MOPS gel by visualizing under UV light in gel documentation unit, Gbox3 (SYNGENE G: Box, USA). RNA was converted to cDNA using 1st strand cDNA synthesis kit (TakaraPrimeScript^{TM} Takara, Japan) as per manufactures' instruction. The confirmation of cDNA was done through polymerase chain reaction (PCR) amplification with 26S rRNA as an internal control (CACAATGATAGGAGGAGCCGAC and CAAGGGAACGGGCTTGGCAGAATC).

PCR primer design and quantitative real-time PCR

qRT PCR primers were designed for selected genes using Primer3 software (Thornton and Basu 2010). qRT-PCR was carried out using SYBER Green™ Premix Ex Taq (Promega) in LightCycler96 Real-Time PCR System (Roche Applied Science, Germany). The reaction mixture contained 10µl of 2X PCR SYBR green ready mix, 1µl of each primer and the cDNA template (2µl) in a final volume of 20µl at 94^{o}C for 3 min followed by 40 cycles of 94^{o}C for 10 sec, 60^{o}C for 30 sec, 55^{o}C for 30 sec. For each sample, the transcript abundance of potential candidate genes was analyzed across a series of three biological replicates for each developmental stage. *TaActin* gene primer was used as an internal control to normalize gene expression (Guan et al. 2019). Melting curve analysis was performed at 55□C with the help of the LightCycler96 software package supplied by Roche. The C_q quantification cycle values for both the reference and target genes were estimated in each sample. The 4DPA stage was considered the control stage for analyzing expression levels of genes at other developing stages. The ΔC_q is the normalized value, and $\Delta\Delta C_q$ is the actual difference of gene expression between control and other developing stages. The

$2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression level in terms of fold change (Livak and Schmittgen 2001). Statistical data analysis was done according to the $\Delta\Delta C_q$ method based on relative expression and fold change values:

$\Delta C_q = \Delta C_{q\ (target)} - \Delta C_{q\ (reference)}$

$\Delta\Delta C_q = \Delta C_{q\ test\ sample\ (4DPA)} - \Delta C_{q\ calibrator\ sample\ (other\ developmental\ stages)}$

Fold change = $2^{-\Delta\Delta CT}$

Results

Wheat grain yield has improved exponentially since the green revolution and continues to improve, although the pace of increment is decreasing. Despite achieving a reasonable level of global production of wheat, work for improving yields has always remained at the forefront, one to fulfill demands of ever-increasing mouths and other to answer the scientific curiosity of “how much yield potential can be improved?” Wild species of wheat encompass a wide diversity of alleles, and *T. dicoccoides* is one of the important species, housing a vast variation in grain size and novel alleles for grain size-related component traits (Gegas et al. 2010; Nevo 2001). For allowing more targeted selection in wheat breeding, the characterization of grain yield components ought to be exploited (Wurschum et al. 2018). TGW is one of the incredibly important yield components in hexaploid wheat, composed of different individual grain morphometric components like GL, GW, and GA (Fuller 2007; Kuchel et al. 2007). For any improvement in TGW, insight into grain size component traits is required to achieve the targeted improvement. In the present study, we used *T. dicoccoides* accessions showing variations in grain size and weight to understand the effect of different genes on the GL, GW, and GA during the grain development from 4DPA to 28DPA and maturity. The selected *T. dicoccoides* accessions included two accessions with higher TGW (33.74g and 31.72g) and the other two with low TGW (17.83g and 14.58g), with slight variations in grain size parameters.

In this study, we tried to gain insight into the effect of multiple genes during grain development across the wheat spikes and within the individual grains. Focusing on specific grain morphometric components at different grain development stages will provide a genetic dissection of TGW along with the mechanistic understanding of genes involved in it and their level of expression. This knowledge plays a vital role in modulating grain yield components to ensure the improvement in wheat yield by understanding the molecular mechanism underlying (Guan et al. 2019; Yu et al. 2019).

In wheat, anthesis begins in the central part of the spike and continues bidirectionally towards the basal and apical parts. Furthermore, the proximal/primary florets of the central spikelet are fertilized two to four days earlier than the distal florets; hence grains from these florets usually have higher weight (Bonnett 1936; Kirby 1974; Peterson 1965; Simmons and Crookston 1979). In the present study, the primary floret of the central three spikelets was selected to assess the different grain parameters to maintain uniformity in the experimental material.

The grain development in wheat has been divided into three phases after anthesis, grain enlargement (0-14DPA), grain filling (15-35DPA), and physiological maturity (36-50DPA). There is a significant increase observed in the length of the developing grain after an initial period of isotropic growth, and this would become maximum at around 15DPA, which contributed more towards grain area, compared to width (Brinton et al. 2017; Xie et al. 2015), suggests that initially effect of GL is more on TGW (Hasan et al. 2011; Lizana et al. 2010). After this, endosperm expands, and grain filling starts with a higher rate at 14-28DPA (Shewry et al. 2012). In the present study, a similar pattern of the rate of increase of GL, GW, and GA was observed. We observed that the increase in GL and GA in the tetraploid *T. dicoccoides* is more prominent between 4DPA to 8DPA (5.44-8.92mm in LG and 4.50-7.33mm in SG accessions), while the maximum gain in GW is between 8DPA to 12DPA (2.10-3.05mm in LG and 1.71-2.48mm in SG accessions). After that, there is a gradual increase in GL, GW, and GA, with the rate of gain decreasing from 16DPA to 28DPA (Brinton et al. 2017). The temporal differentiation of developmental stages of the selected lines from the previous studies may indicate the differences in the genetic background of modern-day bread wheat and wild emmer wheat

Grain size is a complex trait with multiple subcomponents under independent genetic control (Brinton and Uauy 2019; Gegas et al. 2010), and grain development plays an important role in the final TGW. The transition between grain formation and maturation in cereal seed development involves a number of genes that initiate seed size development and decide the ultimate grain yield of wheat. We tried to identify the molecular mechanism behind this process by expression analysis of genes involved in grain size increment (Fig. 4; Supplementary fig S1 and S2). In the present study, we shortlisted eight genes involved in determining different grain size components in rice and wheat. *GL7* (Wang et al. 2015b), *TaGL3* (Qi et al. 2012; Yang et al. 2019), *SRS3* (Kitagawa et al. 2010; Si et al. 2016; Yu et al. 2019), and *TAGASR7* (Huang et al. 2012; Zhang et al. 2014) are well known to affect the grain length, *GS3* (Fan et al. 2006) and *TGW6* (Hanif et al. 2016; Ishimaru et al. 2013) for grain length and weight, *TaGS5* (Ma et al. 2016) for grain size, and *TaGW2* (Simmonds et al. 2016; Su et al. 2011; Wang et al. 2018; Yang et al. 2012; Zhang et al. 2018) is known to affect grain width and weight. The orthologue identification of these genes in *T. dicoccoides* and validation of their expression profiles will open up a new avenue of novel genetic resources to improve the grain size and weight.

GL7 in rice is a major QTL for grain length that controls the grain size and shape through cell elongation and by decreasing cell expansion in terms of grain width direction to produce longer grains (Wang et al. 2015a; Wang et al. 2015b; Zhou et al. 2015). We observed higher expression of *GL7* in LG accessions of *T. dicoccoides* than short grain accessions, confirming that this gene is expressed to produce longer and heavier grains. Due to the increased copy number or mutations in the promoter, the expression of *GL7* was higher in long grain varieties of rice (Wang et al. 2015b). The sequence analysis of this gene in long grain accessions of *T. dicoccoides* might lead to the identification of novel alleles associated with longer grains.

GS5 functions as a positive regulator of grain size and higher expression of *GS5* is correlated with larger grain size (Li et al. 2011; Xu et al. 2015). We observed higher expression levels of *TaGS5* at initial grain development stages in long grain accessions, whereas very low levels in all the developing stages of short grain accessions, indicating that higher expression of *TaGS5* might be involved in the development of larger

grains. Ma et al. (2016) investigated the temporal and spatial expression patterns of the *TaGS5* homoeologues ortholog of rice gene *OsGS5* in various tissues of wheat, which showed higher expression in seedlings, young spikes, and developing grains.

GS3 in rice encodes a putative transmembrane protein, and a major QTL for grain length, weight, and a minor QTL for grain width have been identified. A loss of function of allele in *GS3* promotes cell proliferation and forms long grains, while gain of function produces short grains (Fan et al. 2006). The cause of mutation by premature stop-codons between grain size in rice suggests that orthologous genes and similar related regulatory processes for this type of traits may be conserved across a broad range of taxa ranging from monocot to dicot species. Our result showed higher expression of *GS3* in mid-grain developmental stages in long grain accessions, indicating the role of this orthologue in regulating the grain length in wheat.

In rice, *GL3* encodes a protein phosphatase with kletch-like repeat domains (*OsPPKLs*), restricting cell division in spikelet hulls that increase grain length, weight, and yield (Qi et al. 2012; Zhang et al. 2012). Our results showed higher expression of *TaGL3* in both the LG accession at an early stage of grain development, whereas a low expression level was observed in all the stages of grain development in SG accession. Yang et al. (2019) cloned a wheat orthologous *TaGL3-5A,* its expression pattern was similar with increasing grain size at the early (8DPA) and middle stages (20DPA) of seed development, suggesting that *TaGL3* play a role at an early phase of seed development. Association analysis revealed that the *TaGL3-5A-G* allele was significantly correlated with longer grains and higher TGW, and the frequency of *the* allele in hexaploid wheat was slightly lower than in *T. dicoccoides.*

SRS3, a kinesin 13 protein family gene, regulates seed length in rice and produces long grains by cell elongation (Kitagawa et al. 2010; Si et al. 2016). Yu et al. (2019) conducted a transcriptome profile study in wheat to unravel the genetic architecture of grain size, where the homolog of rice *SRS* showed successively higher expression across the early to middle stages of grain development. Our study also observed higher expression of this gene at an early and middle stage of grain development, which correlates with the phenotypic gain of grain size in LG accessions.

OsGASR7 in rice showed similarity to *Arabidopsis GASA4* and was considered as a candidate gene determining grain length (Huang et al. 2012). The wheat ortholog of *OsGASR* was also reported to play the same role and was involved in grain length development (Dong et al. 2014). Zhang et al. (2014) studied expression patterns of *TaGASR7* in immature seeds in a synthetic hexaploid wheat. *GASR7B* was highly expressed from 6 DPA till 14 DPA in developing seeds of the tetraploid accessions and began to decrease at 17 DPA. Our results also showed an increased expression till 12DPA, indicating that *TaGASR7* was also involved in grain length increase during seed development in *T. dicoccoides*.

GW2 in rice and *Arabidopsis* affect grain size by suppressing cell proliferation (Song et al. 2007; Xia et al. 2013). Our result showed a negative correlation between *TaGW2* and long grain in *T. dicoccoides*, as this gene was highly expressed in SG accessions than large ones. *TaGW2* has been associated with kernel width and weight, which has been validated as a negative regulator of grain size in wheat by gene editing

and mutant analysis. The association analysis indicated that the mutated *TaGAW2* allele significantly increased kernel width (KW) and thousand-kernel weight (TKW) and slightly improved kernel length (KL) in tetraploid and hexaploid wheat. The increase in grain width and length was consistent across grains of different sizes, suggesting that the effect of the mutation is stable across the ear and within spikelets (Simmonds et al. 2016; Su et al. 2011; Wang et al. 2018; Yang et al. 2012; Zhang et al. 2018).

Ishimaru et al. (2013) identified a novel gene for grain length, and weight, *TGW6,* which encodes a novel protein related to indole-3-acetic acid (IAA) synthesis, loss of function of this allele enhances grain length and weight. *TaTGW6-A1*, an ortholog of rice *TGW6*, is associated with grain weight and yield in bread wheat (Hanif et al. 2016). Very low expression of *TGW6* was found in all the accessions used in the present study across different grain developmental stages, indicating a functional allele in the selected *T. dicoccoides* accessions.

Notably, we observed the expression peak of six genes in long grain accession at around 8DPA-20DPA, corresponding to the change in grain size, suggesting that these genes play an important role in the early phase of grain development. There was also a correlation observed between gene expression and grain size traits at the early and middle stages during seed development. Cluster analysis illustrated by the heat map showed that the expression pattern of *GL7*, *TaGL3*, *TaGS5*, *GS3*, *SRS3*, and *TaGASR7* has more similarity to grain size traits (Fig. 5). In contrast, *TaGW2 and TGW6* showed a negative correlation with grain size traits in developing stages.

As there is a correlation between grain size-related traits and expression profiles of selected genes at the early stages of grain development, it depicts the scope of targeting the grain development process to improve yield. *T. dicoccoides* has a large variation in grain size; thus, we speculate that allelic variations of multiple genes involved in grain size are responsible for grain size variation in *T. dicoccoides*. Although other studies have reported these genes in wheat, but expression patterns of these genes in the present study indicate the presence of novel alleles in the *T. dicoccoides* germplasm. The long grain *T. dicoccoides* accessions studied in the present investigation and characterized genes can form a basis for systematic marker-assisted breeding for enhancing the grain size of the breeder's germplasm.

References

1. Bednarek J, Boulaflous A, Girousse C, Ravel C, Tassy C, Barret P, Bitonti B (2012) Down-regulation of the *TaGW2* gene by RNA interference results in decreased grain size and weight in wheat methylation and chromatin patterning. J Exp Bot 63:5945–5955
2. Bonnett O (1936) The development of the wheat spike. J Agric Res 53:445–451
3. Borrill P, Ramirez-gonzalez R, and Uauy C (2016) expVIP: a Customizable RNA-seq data analysis and visualisation platform. Plant Physiol 170:2172–2186
4. Breseghello F, Sorrells ME (2007) QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. Field Crop Res 101(2): 172–179

5. Brinton J, Uauy C (2019) A reductionist approach to dissecting grain weight and yield in wheat. J Intgr Plant Biol 61:337–358
6. 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
7. Brinton J, Simmonds J, Uauy C (2018) Ubiquitin-related genes are differentially expressed in isogenic lines contrasting for pericarp cell size and grain weight in hexaploid wheat. BMC Plant Biol 18:1–17
8. Clavijo BJ, Venturini L, Schudoma C, Accinelli GG, Kaithakottil G, Wright J, Borrill P, Kettleborough G, Heavens D, Chapman H, Lipscombe J, Barker T, Lu FH, McKenzie N, Raats D, Ramirez-Gonzalez RH, Coince A, Peel N, Percival-Alwyn L, Duncan O, Trösch J, Yu G, Bolser DM, Namaati G, Kerhornou A, Spannagl M, Gundlach H, Haberer G, Davey RP, Fosker C, Palma FD, Phillips AL, Millar AH, Kersey PJ, Uauy C, Krasileva KV, Swarbreck D, Bevan MW, Clark MD (2017) 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 5:885–896.
9. Dong LL, Wang FM, Liu T, Dong ZY, Li AL, Jing RL, Mao L, Li YW, Liu X, Zhang KP, Wang DW (2014) Natural variation of *TaGASR7-A1* affects grain length in common wheat under multiple cultivation conditions. Mol Breeding 34:937-947
10. Duan P, Ni S, Wang J, Zhang B, Xu R, Wang Y, Chen H, Zhu X, Li Y (2015) Regulation of *OsGRF4* by *OsmiR396* controls gain size and yield in rice. Nat Plants 2:15203. https://doi.org/10.1038/nplants.2015.203
11. Eversole K, Feuillet C, Mayer KF, Rogers J (2014) Slicing the wheat genome. Science 345:285–287
12. Fan C, Xing Y, Mao H. (2006) 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 112:1164–1171
13. FAO, IFAD, WFP (2015) The state of food insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome, Italy: FAO.
14. Feldman M, Kislev ME (2007) Domestication of emmer wheat and evolution of free-threshing tetraploid wheat. Israel J Plant Sci 55:207–221
15. Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. Ann Bot 100:903–924
16. Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape variation in wheat. The Plant Cell 22:1046–1056
17. Guan Y, Li G, Chu Z, Ru Z, Jiang X, Wen Z, Zhang G, Wang Y, Zhang Y, Wei W (2019) Transcriptome analysis reveals important candidate genes involved in grain-size formation at the stage of grain enlargement in common wheat cultivar “Bainong 4199”. PLos ONE 14(3):e0214149.
18. Hanif M, Gao F, Liu J, Wen W, Zhang Y, Rasheed A, Xia X, He Z, Cao S (2016) *TaTGW6 - A1*, an ortholog of rice *TGW6*, is associated with grain weight and yield in bread wheat. Mol Breed 36:1-8

19. Hasan AK, Herrera J, Lizana C, Calderini DF (2011) Carpel weight, grain length and stabilized grain water content are physiological drivers of grain weight determination of wheat. Field Crops Res 123:241–247
20. Hu J, Wang Y, Fang Y, Zeng L, Xu J, Yu H, Shi Z, Pan J, Zhang D, Kang S (2015) A rare allele of *GS2* enhances grain size and grain yield in rice. Mol Plant 5:1455-1465
21. Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu C, Li J, Fu X (2009) Natural variation at the *DEP1* locus enhances grain yield in rice. Nat Genet 41:494–497
22. Huang XH, Zhao Y, Wei XH, Li CY, Wang A, Zhao Q, Li WJ, Guo YL, Deng LW, Zhu CR, Fan DL, Lu YQ, Weng QJ, Liu KY, Zhou TY, Jing YF, Si LZ, Dong GJ, Huang T, Lu TT, Feng Q, Qian Q, Li JY, Han B (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44:32-39
23. IWGSC (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 6403(361):eaar7191 https://doi.org/10.1126/science.aar7191
24. Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N,Onodera H, Kashiwagi T, Ujiie K, Shimizu B, Onishi A, Miyagawa H, Katoh E (2013) Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. Nat Genet 45:707-711
25. Kassambara A, Kassambara MA (2020) Ggpubr: 'ggplot2' based publication ready plots
26. Kirby E (1974) Ear development in spring wheat. J Agric Sci 82:437–447
27. Kitagawa K, Kurinami S, Oki K, Abe Y, Ando T, Kono I, Yano M, Kitano H, Iwasaki Y (2010) A novel kinesin 13 protein regulating rice seed length:The causal gene of a novel small and round seed mutant. Plant Cell Physiol 51(8):1315–1329
28. Kuchel H, Williams KJ, Jeffries PS (2007) Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World. Theor Appl Genet 115:1029–1041
29. Kumar A, Mantovani EE, Seetan R, Soltani A, Echeverry-Solarte M, Jain S, Simsek S, Doehlert D, Alamri MS, Elias EM, Kianian SF, Mergoum M (2016) Dissection of genetic factors underlying wheat kernel shape and size in an elite x nonadapted cross using a high-density SNP linkage map. Plant Genome 9:2–22
30. Li N, Li Y (2016) Signaling pathways of seed size control in plants. Curr Opin Plant Biol 33:23–32
31. Li W, Yang B (2017) Translational genomics of grain size regulation in wheat. Theor Appl Genet 6:7-14
32. Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nat Genet 6:1-5
33. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\delta\delta CT}$. methods 408:402–408
34. 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.). J Exp Bot 61:1147–1157

35. Ma L, Li T, Hao C, Wang Y, Chen X, Zhang X (2016) *TaGS5-3A*, a grain size gene selected during wheat improvement for larger kernel and yield. Plant Biotech J 14:1269–1280
36. Mao H, Sun S, Yao J, Wang C, Yu S, Xu C, Li X, Zhang O (2010) Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc Natl Acad Sci USA 107:19579-19584
37. Nevo, E., 2001. Evolution of genome-phenome diversity under environmental stress. Proc. Natl. Acad. Sci. U. S. A. 98:6233–6240. https://doi.org/10.1073/pnas.101109298
38. Pearce S, Vazquez-Gross H, Herin SY, Hane D, Wang Y, Gu YQ, Dubcovsky J (2015) WheatExp: an RNA-seq expression database for polyploid wheat. BMC Plant Biol 15:299. https://doi.org/10.1186/s12870-015-0692-1
39. Peterson RF (1965) Wheat: botany, cultivation and utilisation. London, Leonard Hill. 448 pp.
40. Qi P, Lin YS, Song XJ, Shen JB, Huang W, Shan JX, Zhu MZ, Jiang L, Gao JP, Lin HX (2012) The novel quantitative trait locus *GL3.1* controls rice grain size and yield by regulating Cyclin-T1;3. Cell Res 22:1666-1680
41. R Core Team 2018 R:A language and environment for statistical computing. Vienna, Austria. URL http://www.R-project.org/.
42. Ray DK, Mueller ND, West PC, Foley JA (2013) Yield trends are insufficient to double global crop production by 2050. PLoS ONE 8:e66428
43. Segami S, Kono I, Ando T, Yano M, Kitano H, Miura K (2012) Small and round seed 5 gene encodes alpha-tubulin regulating seed cell elongation in rice. Rice 5:4
44. Shewry PR, Mitchell RAC, Tosi P, Wan Y, Underwood C, Lovegrove A, Freeman J, Toole GA, Mills ENC, Ward JL (2012) An integrated study of grain development of wheat (cv. Hereward). J Cereal Sci 56:21–30
45. Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S (2008). Deletion in a gene associated with grain size increased yields during rice domestication. Nature 40:1024-28
46. Si L, Chen J, Huang X, Gong H, Luo J, Hou Q, Zhou T, Lu T, Zhu J,Shangguan Y, Chen E, Gong C, Zhao Q, Jing Y, Zhao Y, Li Y, Cui L, Fan D, Lu Y, Weng Q, Wang Y, Zhan Q, Liu K, Wei X, An K, An G, Han B (2016) *OsSPL13* controls grain size in cultivated rice. Nat Genet 48:447-456
47. 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 Biol 14:1–13
48. Simmons SR, Crookston RK (1979) Rate and duration of growth of kernels formed at specific florets in spikelets of spring wheat. Crop Sci 19: 690-693.
49. Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, Jorge B, Uauy C (2016) A splice acceptor site mutation in *TaGW2 - A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor Appl Genet 129:1099–1112
50. Song X J, Huang W, Shi M, Zhu M-Z, Lin H X (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39:623–630

51. Su Z, Hao C, Wang L, Dong Y, Zhang X (2011) Identification and development of a functional marker of *TaGW2* associated with grain weight in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 122:211-223
52. Wan Y, Poole RL, Huttly AK, Toscano-Underwood C, Feeney K, Welham S, Gooding MJ, Mills C, Edwards KJ, Shewry PR, Mitchell R (2008) Transcriptome analysis of grain development in hexaploid wheat. BMC Genomics 9:121. https://doi.org/10.1186/1471-2164-9-121
53. Wang S, Li S, Liu Q, Wu K, Zhang J, Wang Y, Chen X, Zhang Y, Gao C, Wang F, Huang H, Fu X (2015a) The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nat Genet 47:949-954
54. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J; International Wheat Genome Sequencing Consortium, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo MC, Dvorak J, Morell M, Dubcovsky J, Ganal M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, Hayden M, Akhunov E (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotech J 12:787–796
55. Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, Zhang G, Fu X (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet 44:950-954
56. Wang W, Simmonds J, Pan Q, Davidson D, He F, Battal A, Akhunova A, Trick HN, Uauy C, Akhunov E (2018) Gene editing and mutagenesis reveal inter - cultivar differences and additivity in the contribution of *TaGW2* homoeologues to grain size and weight in wheat. Theor Appl Genet 131:2463–2475
57. Wang Y, Xiong G, Hu J, Jiang L, Yu H, Xu J, Fang Y, Zeng L, Xu E, Ye W, Meng X, Liu R, Chen H, Jing Y, Wang Y, Zhu X, Li J, Qian Q (2015b) Copy number variation at the *GL7* locus contributes to grain size diversity in rice. Nat Genet 47:944-948
58. Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X (2008) Isolation and initial characterization of *GW5,* a major QTL associated with rice grain width and weight. Cell Res 18:1199-1209
59. Whan AP, Smith AB, Cavanagh CR, Ral JF, Shaw LM, Howitt CA (2014) GrainScan : a low cost, fast method for grain size and colour measurements. Plant Methods 10:23. https://doi.org/10.1186/1746-4811-10-23
60. Wickham H (2016) GGplot2: Elegant graphics for data analysis. https://ggplot2.tidyverse.org
61. Winfield MO, Allen AM, Burridge AJ, Barker GLA, Benbow HR, Paul A, Coghill J, Waterfall C, Davassi A, Scopes G, Pirani A, Webster T, Brew F, Bloor C, King J, West C, Griffiths S, King I, Bentley AR, Edwards KJ (2016) High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotech J 796:1195–1206
62. Wurschum T, Leiser WL, Langer SM, Tucker MR, Longin CFH (2018) Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long - term genetic trends of grain yield components. Theor Appl Genet 131:2071–2084
63. Xia T, Li N, Dumenil J, Li J, Kamenski A, Bevan MW, Gao F, Li Y (2013) The ubiquitin receptor DA1 interacts with the E3 ubiquitin ligase DA2 to regulate seed and organ size in Arabidopsis. The Plant Cell

Online 25:3347–3359

64. Xie Q, Mayes S, Sparkes DL (2015) Carpel size, grain filling, and morphology determine individual grain weight in wheat. J Exp Bot 66:6715–6730
65. Xu C, Liu Y, Li Y, Xu X, Xu C, Li X, Xiao J, Zhang Q (2015) Differential expression of *GS5* regulates grain size in rice. J Exp Bot 66:2611–2623
66. Yang J, Zhou Y, Wu Q, Chen Y, Zhang P, Zhang Y, Hu W, Wang X, Zhao H, Dong L, Han J, Liu Z, Cao T (2019) Molecular characterization of a novel TaGL3-5A allele and its association with grain length in wheat (*Triticum aestivum* L.). Theor Appl Genet 132:1799–1814
67. Yang Z, Bai Z, Li X, Wang P, Wu Q, Yang L, Li L, Li X (2012) SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. Theor Appl Genet 125:1057–1068
68. Yu K, Liu D, Chen Y, Wang D, Yang W, Yang W, Yin L (2019) Unraveling the genetic architecture of grain size in einkorn wheat through linkage and homology mapping and transcriptomic profiling. Theor Appl Genet 70:4671-4687
69. Zhang D, Wang B, Zhao J, Zhao X, Zhang L, Liu D, Dong L, Wang D, Mao L, Li A (2014) Divergence in homoeolog expression of the grain length-associated gene GASR7 during wheat allohexaploidization. The Crop J 1:1–3 https://doi.org/10.1016/j.cj.2014.08.005
70. Zhang G, Wang Y, Guo Y, Zhao Y, Kong F, Li S (2015) Characterization and mapping of QTLs on chromosome 2D for grain size and yield traits using a mutant line induced by EMS in wheat. Crop J 3:135–144
71. Zhang Y, Liu J, Xia X (2014) TaGS - D1, an ortholog of rice OsGS3, is associated with grain weight and grain length in common wheat. Mol Breeding 34:1097-1107
72. Zhang Y, Li D, Zhang D, Zhao X, Cao X, Dong L, Liu J, Chen K (2018) Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits. The Plant J 94:857-866
73. Zhang X, Wang J, Huang J, Lan H, Wang C, Yin C, Wu Y, Tang H, Qian Q, Li J, Zhang H (2012) Rare allele of *OsPPKL1* associated with grain length causes extra-LG and a significant yield increase in rice. Proc Natl Acad Sci USA 109:21534- 21539
74. Zhou Y, Miao J, Gu H, Peng X, Leburu M, Yuan F, Gu H, Gao Y, Tao Y, Zhu J, Gong Z, Yi C, Gu M, Yang Z, Liang G (2015) Natural variations in *SLG7* regulate grain shape in rice. Genetics 201:1591-1599
75. Zuo J, Li J (2014) Molecular genetic dissection of quantitative trait loci regulating rice grain size. Annu Rev Genet 48:99–118

Tables

Table 1 Mean values (Mean ± S.E.) recorded by the *T. dicoccoides* accessions for grain yield-related traits

	pau5228	pau5232	pau14703	pau14756
GL	10.17 ± 0.06^{a}	09.98 ± 0.07^{a}	08.3 ± 0.12^{b}	08.09 ± 0.15^{b}
GW	02.90 ± 0.15^{a}	02.98 ± 0.06^{a}	02.79 ± 0.10^{a}	01.91 ± 0.07^{b}
GA	23.02 ± 0.85^{a}	21.87 ± 0.59^{a}	17.07 ± 0.49^{b}	12.28 ± 0.45^{c}
TGW	33.74 ± 0.11^{a}	31.72 ± 0.04^{b}	17.83 ± 0.03^{c}	14.58 ± 0.10^{d}

Different letters (^{a}, ^{b}, ^{c}, ^{d}) were denoted according to the mean comparison of four *T. dicoccoides* accessions by DMRT analysis at α=0.01%, GL: Grain Length (mm), GW: Grain Width (mm), GA: Grain Area (mm^2), TGW**:** Thousand Grain Weight (g)

Table 2 List of selected genes that are involved in controlling seed size in different crop species

Genes	Identified in Species	Phenotype	Protein category/structure	Function
GL7	Rice	Long grains by overexpression	TON1 RECRUITING MOTIF–containing protein	A major QTL for rice seed length by encouraging cell elongation and repressing cell expansion in favor of seed–width, it regulates grain size
GW2	Rice	Wide grains	a RING-type E3 ubiquitin ligase	underlying a major QTL for rice grain width and weight
qGL3	Rice	Long grains	Protein phosphatase kelch family–Ser/Thr phosphatase (PPKL)	The novel QTL regulating rice seed size and produce by controlling cyclin–T1
GS3	Rice	Long grains	Homolog of G–protein y subunit	a major QTL for seed length and seed width and minor QTL for grain width and thickness. *GS3* forms long seeds because of increased cell proliferation
TGW6	Rice	Long grains	Indole–3–acetic acid IAA–glucose hydrolase	Major QTL down-regulates endosperm development and seed weight
GS5	Rice	Slender grains	Putative serine carboxypeptidase	Natural variation in GS5 Regulates grain size and yield
GASR7	Rice	Longer grains	GSK3/SHAGGY–like regulator of signaling kinase	GSK2, a down regulator of BR signaling, interacts with GS2 and prohibits its transcriptional- activation activity, indicating BR to be involved in GS2–mediated seed size control
SRS3	Rice	Long grain by cell elongation	kinesin 13 protein	Regulates the seed length in rice

Table 3 Primers used for qRT-PCR

Gene	NCBI gene Ds	*T. dicoccoides* Ensemble gene IDs		Primer sequence
GL7	KP89957	TRIDC2AG023260	Forward	TCCTTGACACATCCTTCTACC
			Reverse	CTGGTTTGATCTGGCTGACTTCAC
GS3	Os03g0407400	TRIDC4AG045300	Forward	AAACATGGCAGGGAGGAGAAGG
			Reverse	GCAACAGCTGATTCCTCTTCG
SRS3	AB531488	TRIDC1AG016970	Forward	GCTCACAGGAAGGAAATCGAG
			Reverse	AGCCTTGCGTGACAAAAGAAAG
TaGL3	KY865329	TRIDC5AG054600	Forward	AGTTGCCACAGGGACTGGAT
			Reverse	CGGCCAAAACTTCAAACACA
TaGS5	KX219726	TRIDC3AG023140	Forward	GGATGTTGCCTGGGAAATGG
			Reverse	TGTGGGCTTGTTGAGGGGTA
TaGASR7	KJ000052/53	TRIDC7AG026210	Forward	GGGGACGCAGTACAAGAAGG
			Reverse	CTCCCTCCTTGGTCTTCCAG
TaGW2	JN896622/23	TRIDC6BG062240	Forward	TGCTGGTAGTTAATGACGATGTCC
			Reverse	GTGAGACTAATTTGGAACATACGC
TGW6	Os06g0623700	TRIDC4AG045830	Forward	TTGACCAGAACTACTGTGACTCC
			Reverse	ACTATGCCATCGCAATGGAC

Figures

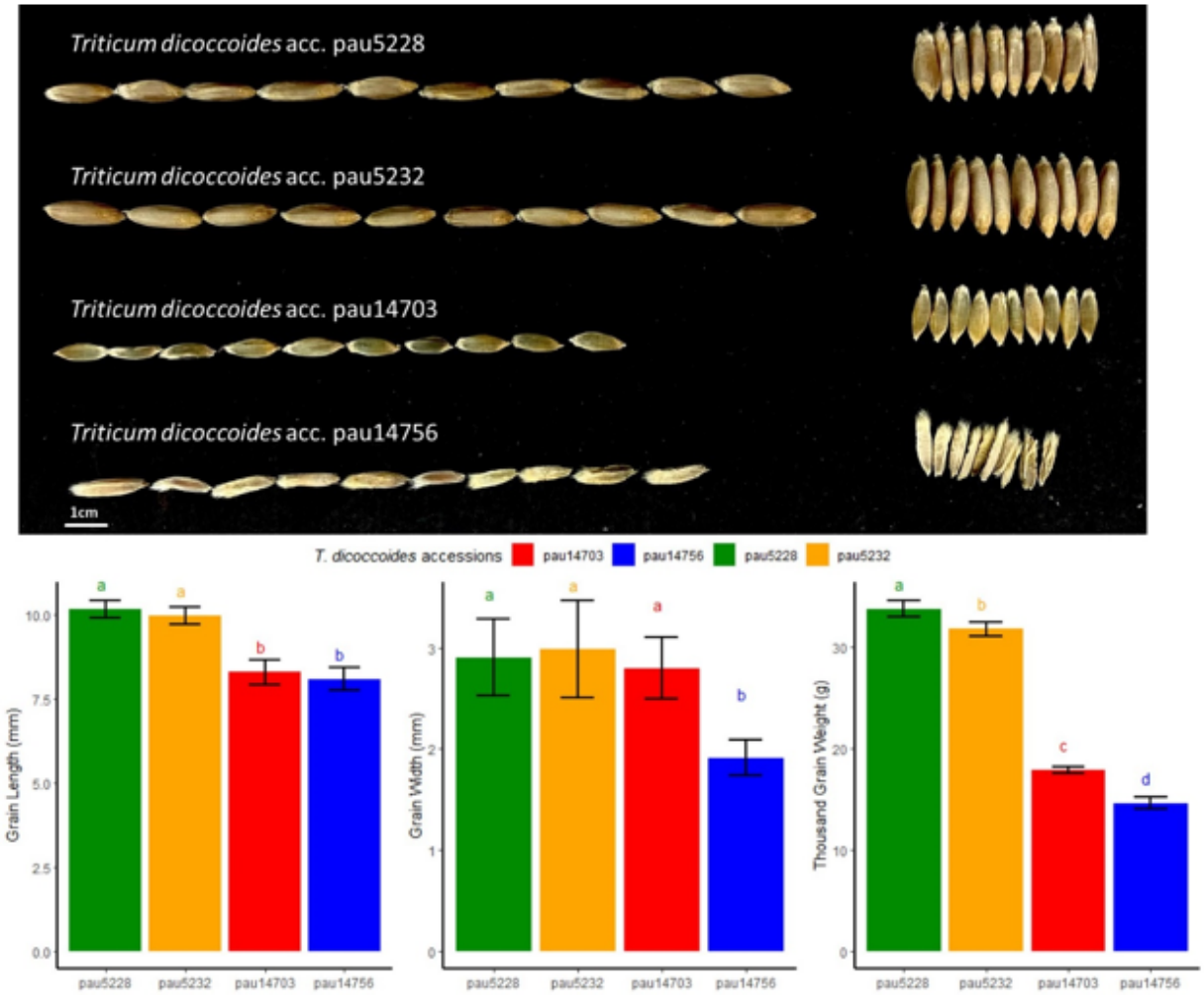

Figure 1

Representative matured grains from four T. dicoccoides accessions. Grains are aligned to show differences in (A) grain length (GL, 10 grains) and (B) grain width (GW, 10 grains). Scale bar = 1 cm (C) Comparison of GL, GW, and Thousand Grain Weight (TGW) in graphical form of T. dicoccoides accessions pau5228, pau5232, pau14703, pau14756. Different small letters indicate significant differences at 0.01% significance level.

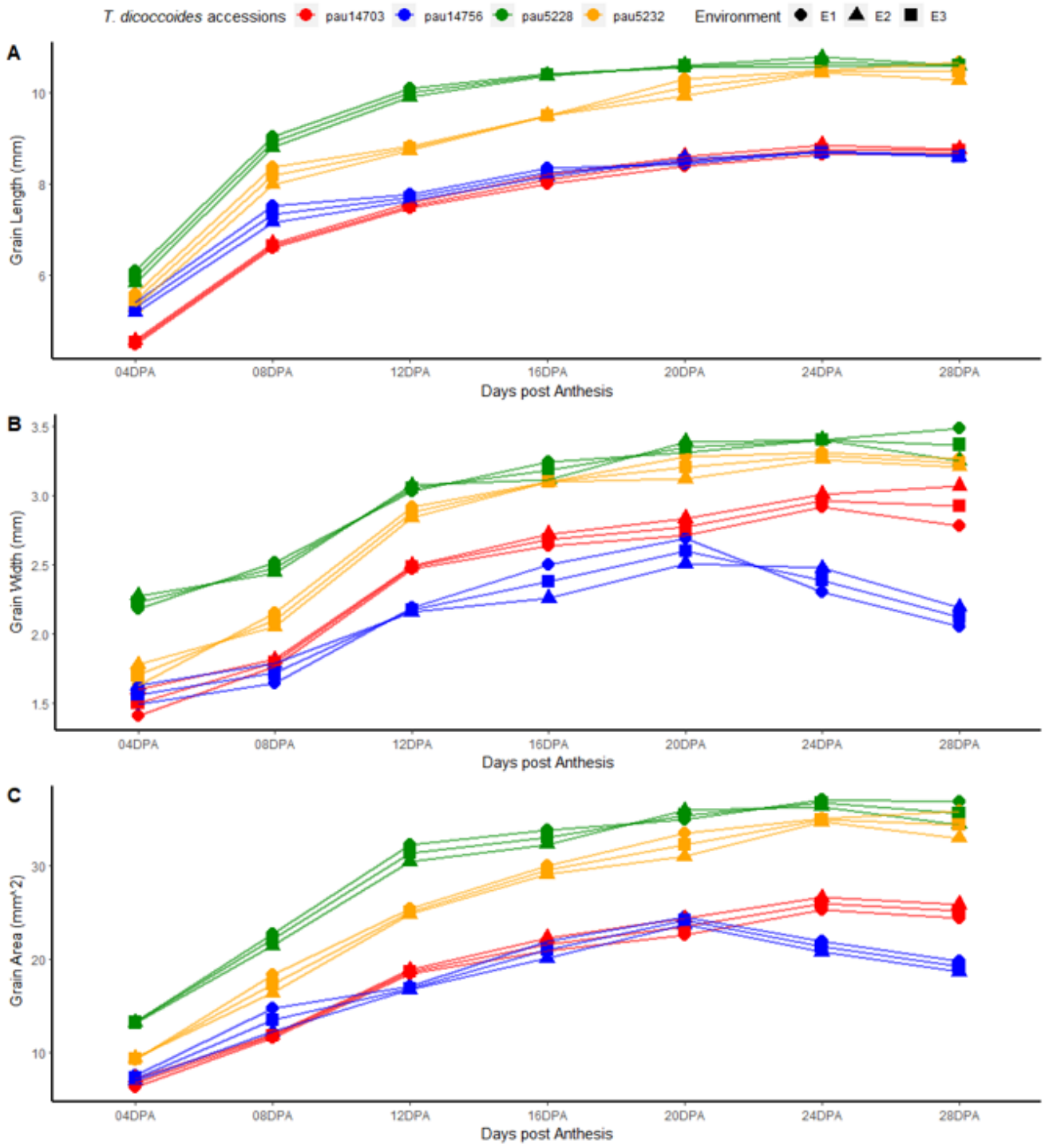

Figure 2

Grain development time course of four T. dicoccoides accessions. (A) Grain length (B) Grain width and (C) Grain area during grain development with the samples taken at 4DPA (days post anthesis), 8DPA, 12DPA, 16DPA, 20DPA, 24DPA, and 28DPA in 2018 (E1), 2019 (E2), and average mean of the two years (E3).

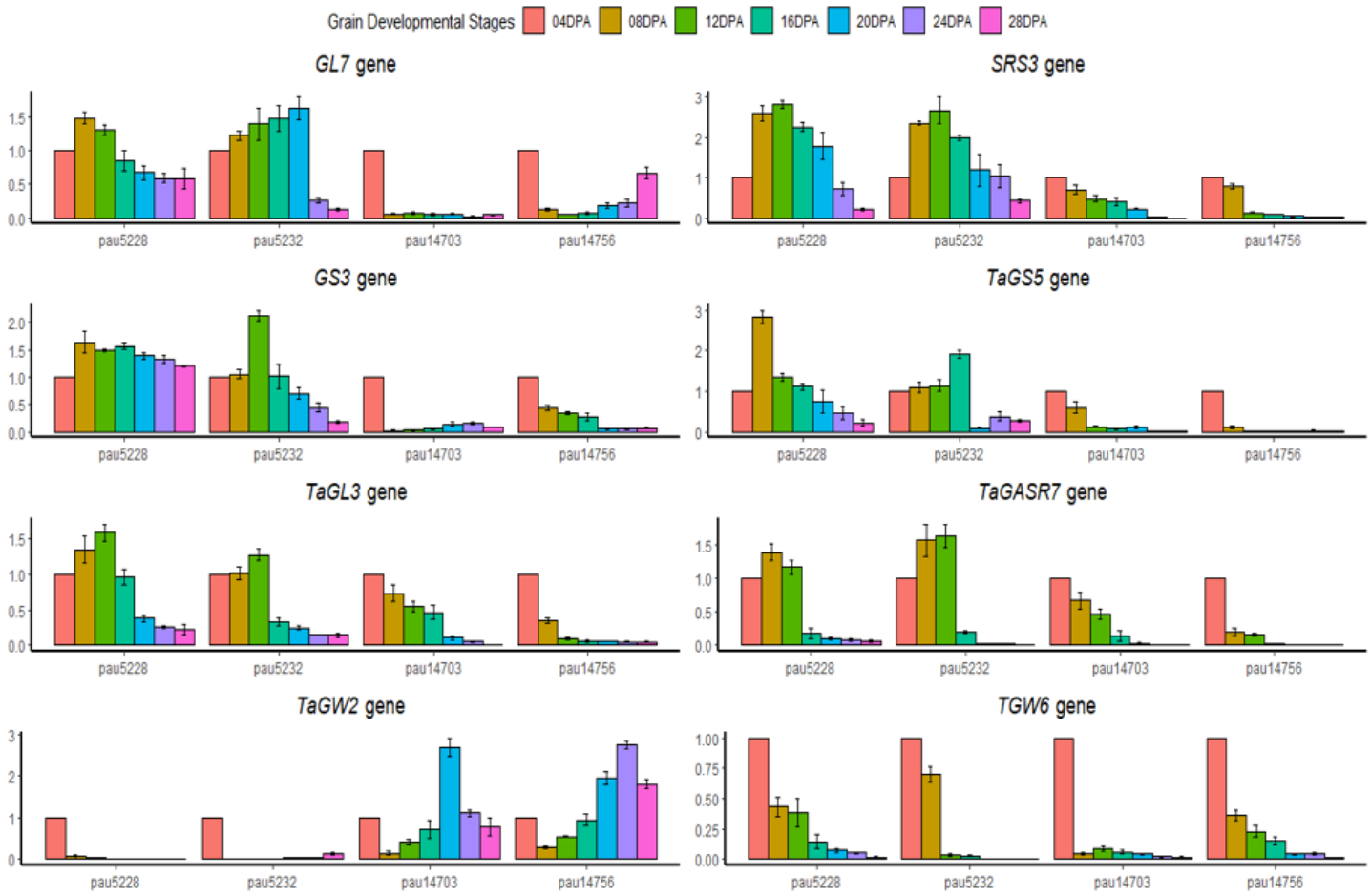

Figure 3

Expression patterns of genes related to grain size formation in four accessions of T. dicoccoides germplasm analyzed by qRT-PCR. Seeds were sampled at 4, 8, 12, 16, 20, 24, and 28DPA (Days post anthesis). The expression of all genes at 4DPA in all accessions was assumed to be 1. Means and standard errors are shown from the analysis of three biological replicates.

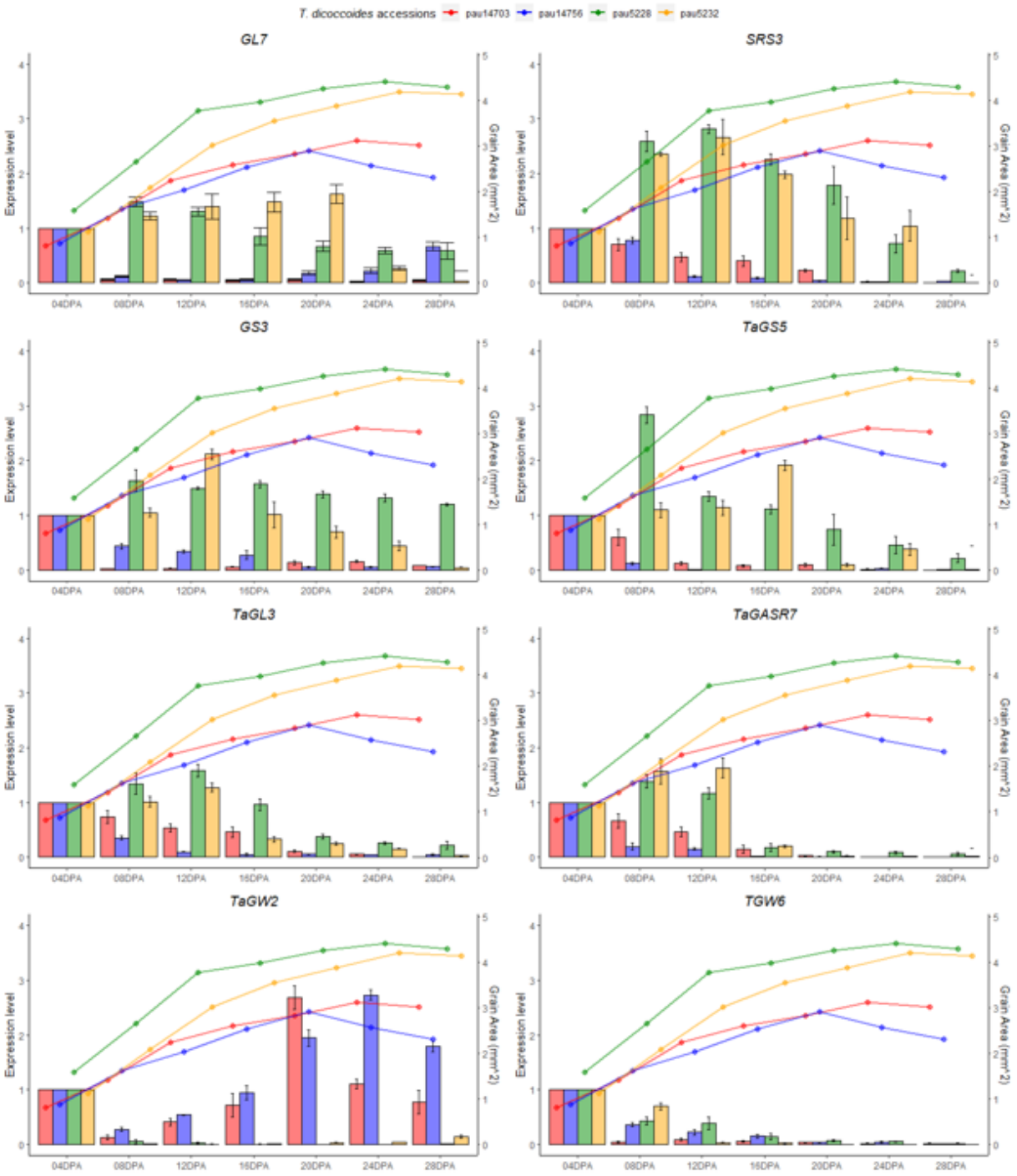

Figure 4

Relation between expression profiles of eight selected genes and phenotypic changes in grain area of T. dicoccoides accessions at different grain developmental stages from 4DPA to 28DPA.

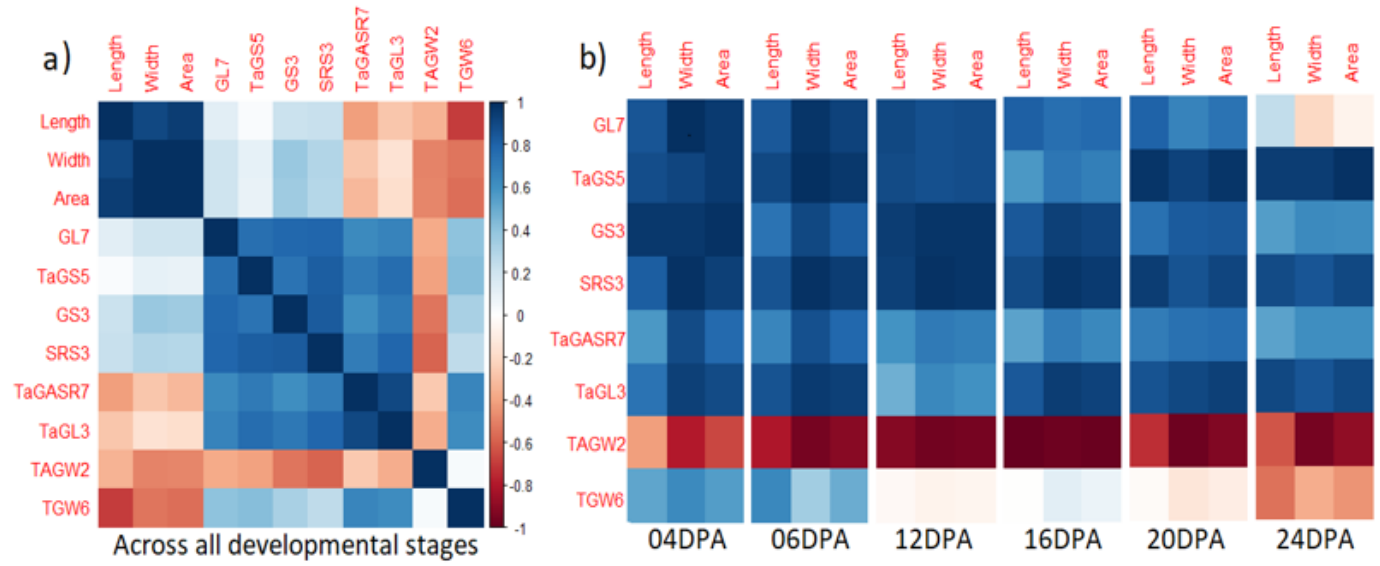

Figure 5

Heat map showing the correlation between (a) expression of all eight genes across all the developmental stages (b) and expression of each gene at each grain developmental stages

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppTable.docx
- SupplFigure.docx